

May2013

Incubator PMC report for May 2013

The incubator PMC continues to have difficulties with supervision and reporting, as evinced by set of missing reports below. [1] It's not really clear why this problem has grown worse. Perhaps the previous chair was more capable of harnessing the efforts of the PMC, or perhaps it's the larger number of podlings, or perhaps it's random, or more likely it's a combination of the above.

While there is a consensus in the PMC that there are problems, there is not a consensus about what actions to take. Several people have proposed changes, and all of those changes have attracted a mixture of support and opposition.

If nothing changes, either at the level of lack of reports and other evidence of supervision problems, or with the PMC's ability to reach a consensus, Benson is going to ask the Board for an agenda item on the June agenda to discuss the situation.

On the positive side, the PMC has added several new members from inside the podlings, which is one of the schemes for getting more supervision done.

[1] There is some evidence that Marvin failed to deliver some reminders. If true, it is a partial explanation but not a complete excuse.

o Community

New IPMC members:

David Nalley
Henry Saputra
Hemming Schmiedenhausen
Chip Childers
Joe Brockmeier

People who left the IPMC:

jeremias@apache.org
Stefan Bodewig

o New Podlings

jclouds

o Graduations: the board has motions for the following:

<none>

o Releases

o Legal / Trademarks

o Infrastructure

The Wiki outage was a minor disruption.

----- Summary of podling reports -----

Still getting started at the Incubator

These projects are still getting started, so no immediate progress towards graduation is yet expected.

Not yet ready to graduate

Ready to graduate

Table of Contents

Ambari
Blur
Curator
DeltaSpike
Droids
Falcon
Hadoop Development Tools
jclouds
Knox
Mesos
MRQL
NPanday
ODF Toolkit
Onami
Open Climate Workbench
Provisionr
Tajo
Tez

Ambari

Ambari is a monitoring, administration and lifecycle management project for Apache Hadoop clusters.

Ambari has been incubating since 2011-08-30.

Three most important issues to address in the move towards graduation:

- 1.
- 2.
- 3.

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

How has the community developed since the last report?

How has the project developed since the last report?

Please check this [] when you have filled in the report for Ambari.

Signed-off-by:

Owen O'Malley: [](ambari)

Chris Douglas: [](ambari)

Arun Murthy: [](ambari)

Shepherd notes:

Blur

Blur is a search platform capable of searching massive amounts of data
in a cloud computing environment.

Blur has been incubating since 2012-07-24.

Three most important issues to address in the move towards graduation:

1. An official Apache Release - currently have a plan for 0.1.5 being that release.
2. Community growth, Blur still needs to grow it's community.

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

None.

How has the community developed since the last report?

- We've welcomed a new PMC Member/Committee.
- Subscriptions: user@ - 35[+4]; dev@ - 38[+5]

How has the project developed since the last report?

- In pursuing an official release, we settled on solidifying and releasing the current code - putting on pause for now a 0.2 version. This should bring a release much faster.

Please check this when you have filled in the report for Blur.

Signed-off-by:

Doug Cutting: (blur)
Patrick Hunt: (blur)
Tim Williams: (blur)

Shepherd notes:

Curator

Curator - ZooKeeper client wrapper and rich ZooKeeper framework

Curator has been incubating since 2013-03-11.

Three most important issues to address in the move towards graduation:

1. Build community
2. Create a first release
3. Finalize name search/trademarks

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

No.

How has the community developed since the last report?

Outside users are starting to post issues, ask questions, etc.
Several potential new committers have been identified.

There are currently 15 subscribers on the dev list.
There are 18 subscribers to the user list.

How has the project developed since the last report?

The mentors have continued to educate the committers on the Apache Way and how to conduct a proper release.
The initial 2.0.0-incubating release is now very close to finalization.

Please check this when you have filled in the report for Curator.

Signed-off-by:

Enis Söztutar: (curator)
Luciano Resende: (curator)
Mahadev Konar: (curator)
Patrick Hunt: (curator)

Shepherd notes:

DeltaSpike

DeltaSpike is a collection of JSR-299 (CDI) Extensions for building applications on the Java SE and EE platforms.

DeltaSpike has been incubating since 2011-12-07.

Three most important issues to address in the move towards graduation:

- 1.
- 2.
- 3.

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

How has the community developed since the last report?

How has the project developed since the last report?

Please check this when you have filled in the report for DeltaSpike.

Signed-off-by:

Mark Struberg: (deltaspikes)

Gerhard Petracek: (deltaspikes)

David Blevins: (deltaspikes)

Matt Benson: (deltaspikes)

Jim Jagielski: (deltaspikes)

Shepherd notes:

Droids

Droids aims to be an intelligent standalone robot framework that allows to create and extend existing droids (robots).

Droids has been incubating since 2008-10-09.

Three most important issues to address in the move towards graduation:

1. Activity

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

None

How has the community developed since the last report?

There has been no change in the active members of the community.

There are enough active people to form a viable PMC.

How has the project developed since the last report?

Development this quarter has been slow after the flurry of development last quarter. The codebase has a decent level of maturity, so a lot of changes isn't expected.

Please check this when you have filled in the report for Droids.

Signed-off-by:

Thorsten Scherler: (droids)

Richard Frovarp: (droids)

Shepherd notes:

Falcon

Falcon is a data processing and management solution for Hadoop designed for data motion, coordination of data pipelines, lifecycle management, and data discovery. Falcon enables end consumers to quickly onboard their data and its associated processing and management tasks on Hadoop clusters.

Falcon has been incubating since 2013-03-27.

Three most important issues to address in the move towards graduation:

1. Building a site for Falcon
2. Add new and diverse committers
3. Build and grow community
4. Releases at frequent and regular intervals

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

- No

How has the community developed since the last report?

Most of the infrastructure required for the project has been setup now. We are in the process of making it easy for new users to build, try and adopt Falcon. Community is still building.

How has the project developed since the last report?

Few JIRAs have been filed during this period and few users have subscribed to the mailing list.

Please check this [] when you have filled in the report for Falcon.

Signed-off-by:

Arun Murthy: [X](falcon)

Chris Douglas: [X](falcon)

Owen O'Malley: [](falcon)

Devaraj Das: [](falcon)

Alan Gates: [](falcon)

Shepherd notes:

Hadoop Development Tools

Eclipse based tools for developing applications on the Hadoop platform

Hadoop Development Tools has been incubating since 2012-11-09.

Three most important issues to address in the move towards graduation:

- Support multiple versions of Hadoop in a single IDE instance. During this time building understanding of the Apache processes around working and releasing.
- Release
- Grow the podling community in terms of users and contributors.

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

- None

How has the community developed since the last report?

- Mirko Kaempf has begun contributing to the project (HDT-21)
- Worked with, and established a process for, a contribution from a GitHub.com pull request

How has the project developed since the last report?

- Tools now allow interacting with HDFS and launching Map/Reduce jobs on Hadoop 0.23.6
- New class wizards that support the "new" mapreduce api

Please check this [X] when you have filled in the report for Hadoop Development Tools.

Signed-off-by:

Suresh Marru: [X](hadoopdevelopmenttools)

Chris Mattmann: [X](hadoopdevelopmenttools)

Roman Shaposhnik: [](hadoopdevelopmenttools)

Shepherd notes:

jclouds

A cloud agnostic library that enables developers to access a variety of supported cloud providers using one API

jclouds has been incubating since 2013-04-29.

Three most important issues to address in the move towards graduation:

We are in the initial stages of setup. Our mailing lists have just been created, and we are waiting for

infrastructure to migrate git repositories.

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

None

How has the community developed since the last report?

This is our first report.

How has the project developed since the last report?

This is our first report.

Please check this when you have filled in the report for jclouds.

Signed-off-by:

Brian McCallister: [](jclouds)

Tom White: [X](jclouds)

Henning Schmiedehausen: [](jclouds)

David Nalley: [x](jclouds)

Jean-Baptiste Onofré: [](jclouds)

Mohammad Nour El-Din: [X](jclouds)

Olivier Lamy: [x](jclouds)

Tomaz Muraus: [](jclouds)

Suresh Marru: [X](jclouds)

Carlos Sanchez: [](jclouds)

Shepherd notes:

Knox

Knox Gateway is a system that provides a single point of secure access for Apache Hadoop clusters.

Knox has been incubating since 2013-02-22.

Three most important issues to address in the move towards graduation:

1. Expand community to include more diverse committers.
2. Align technically with security work going in in Hadoop.
3. Clear the project name with legal and pick a new name if required.

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

1. None

How has the community developed since the last report?

1. Now that we have a release we will be reaching out to those that expressed interest when the project was proposed.
2. Couple of active developers engaging the project on the mailing lists at the moment.
3. Apache ways of discussing on list and other important knowledge being learned and executed by the community.

How has the project developed since the last report?

1. Completed initial release (0.2.0) via Apache processes.
2. Over 65 issues brought into JIRA.
3. All Knox source code/etc. has been vetted and is in Git.

Please check this [] when you have filled in the report for Knox.

Signed-off-by:

Owen O'Malley: [](knox)

Chris Douglas: [X](knox)

Mahadev Konar: [](knox)

Alan Gates: [X](knox)

Devaraj Das: [](knox)

Chris Mattmann: [X](knox)

Tom White: [X](knox)

Shepherd notes:

Mesos

Mesos is a cluster manager that provides resource sharing and isolation across cluster applications.

Mesos has been incubating since 2010-12-23.

Three most important issues to address in the move towards graduation:

1. Improved documentation and support using ASF Mesos Wiki.
2. Roll an 0.11.0 release.
3. More interaction on the dev lists besides those from Review Board/JIRA.

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

The project currently only has 1 active mentor, Chris Mattmann. Chris has been able to get Paul Ramirez and Andrew Hart, both IPMC members, to help mentor as well when needed (e.g., with release VOTE'ing), but it's still been pretty scarce. Chris brought this up to the IPMC in threads before, but it went without a reply.

How has the community developed since the last report?

Ben Mahler was added as a PPMC member and committer on the project.

We've seen an increase in activity on the mailing list. Airbnb has been increasingly using Mesos, not only for Chronos, but also for Hadoop (using our newly written Hadoop framework). There has also been a production engineering meetup where Mesos was presented:
<http://www.meetup.com/SF-Bay-Area-Large-Scale-Production-Engineering/events/97563092/>

How has the project developed since the last report?

The project has moved its source repo to ASF Git per <https://issues.apache.org/jira/browse/INFRA-6065>.

We need to get our website properly ASF branded. We also need to move towards more discussions on the dev list that aren't shepherded by automated tools (Review Board, and JIRA). This would make it easier to follow conversations, and to encourage new contributors to be able to discuss the code, and its stewardship.

On top of the usual bug fixes and reliability improvements, we've implemented resource monitoring to gain visibility into the resource consumption of executors running in Mesos. We are also much closer to shipping a feature that allows the slave to upgrade gracefully (without killing all tasks underneath it).

Signed-off-by:

Chris Mattmann: [X](mesos)
Brian McCallister [](mesos)
Tom White [X](mesos)

MRQL

MRQL is a query processing and optimization system for large-scale, distributed data analysis, built on top of Apache Hadoop and Hama.

MRQL has been incubating since 2013-03-13.

Three most important issues to address in the move towards graduation:

1. Expand the development community to include more diverse committers
2. Complete the first release
3. Improve query performance on large clusters

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

none

How has the community developed since the last report?

We are in the process of completing the infrastructure for the MRQL podling. We have created a new wiki site for MRQL that includes a detailed developer documentation, which includes a guide to potential committers on how to contribute to the MRQL project, gives a detailed roadmap of the codebase, and lists the future release plans. This wiki also contains user documentation of the MRQL query language. We have also created a JIRA site for MRQL. So far we have created 9 JIRA issues, some of them related to infrastructure improvement, from which 6 have already been resolved.

How has the project developed since the last report?

The existing code has been substantially modified in many ways to facilitate contributions by committers. The codebase now includes a testbed of queries to automatically test various aspects of the system. We are now in the process of refactoring the code based on our coding guidelines and of using maven as our project management tool.

Please check this [] when you have filled in the report for MRQL.

Signed-off-by:

Alex Karasulu: [X](mrql)

Anthony Elder: [](mrql)

Alan Cabrera: [X](mrql)

Mohammad Nour El-Din: [](mrql)

Shepherd notes:

NPanday

NPanday allows projects using the .NET framework to be built with Apache Maven.

NPanday has been incubating since 2010-08-13.

Three most important issues to address in the move towards graduation:

1. work out a concrete plan towards graduation ASAP
2. encourage newer contributors to do so on a continuing basis
3. reach out to other projects using .NET that might be interested in working with NPanday

Some new attention has sparked recently around getting support for Mono back up to par, which might help resolve some of the regressions that are in the way of a release. Hopefully some of the newer contributors will be able to continue helping out in that regard.

Users have drawn attention to the fact that the home page doesn't list any recent activity. Some effort needs to go into making it clearer of the current status and how to contribute.

The existing committers continue to be busy with other projects, but are still interested in seeing the project succeed and looking for new volunteers.

We are still short on mentors, and would appreciate any volunteers.

Please check this [] when you have filled in the report for NPanday.

Signed-off-by:

Dennis Lundberg: [](npanday)

Shepherd notes:

ODF Toolkit

Java modules that allow programmatic creation, scanning and manipulation of OpenDocument Format (ISO/IEC 26300 == ODF) documents

ODF Toolkit has been incubating since 2011-08-01.

Three most important issues to address in the move towards graduation:

1. Grow the community, especially attracting new developers.
2. Successfully complete another podling release.
3. Hold an informal graduation readiness vote.

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

No.

How has the community developed since the last report?

No new committers added.

We stalled in our 0.6 release plans when our Release Manager left due to changing work commitments. A new Release Manager stepped forward and we're currently voting on a new Release Candidate.

There has also been some discussion of whether we will ever become large enough to sustain a TLP, and whether it would make more sense to move the code into an existing TLP, e.g., OpenOffice or POI, both of which are related technologies.

How has the project developed since the last report?

Completed development work for 0.6 release. Along with the usual bug fixes, enhancements in this release include:

- * Added document encryption support
- * Added metadata support
- * Support for OpenDocument-v1.2
- * Additional APIs for Simple API

Please check this [] when you have filled in the report for ODF Toolkit.

Signed-off-by:

Sam Ruby: [](odftoolkit)

Nick Burch: [](odftoolkit)

Yegor Kozlov: [](odftoolkit)

Shepherd notes:

Onami

Apache Onami aims to create a community focused on the development and maintenance of a set of Google Guice extensions not provided out of the box by the library itself nor the Google developers team, such as integration with 3rd part frameworks or extra functionalities.

Onami has been incubating since 2012-11-14.

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

We have successfully performed all votes to graduate. Graduation has not performed, infra issues are still open: <https://issues.apache.org/jira/browse/INFRA-6164>

Please check this [grobmeier] when you have filled in the report for Onami.

Signed-off-by:

Christian Grobmeier: [X](onami)

Mohammad Nour El-Din: [X](onami)

Olivier Lamy: [](onami)

Shepherd notes:

Open Climate Workbench

Apache Open Climate Workbench (Incubating) is an effort to develop software that performs climate model evaluation using model outputs from a variety of different sources (the Earth System Grid Federation, the Coordinated Regional Downscaling Experiment, the U.S. National Climate Assessment and the North American Regional Climate Change Assessment Program) and temporal/spatial scales with remote sensing data from NASA, NOAA and other agencies. The toolkit includes capabilities for regridding, metrics computation and visualization.

Open Climate Workbench has been incubating since 2013-02-15.

Three most important issues to address in the move towards graduation:

1. Develop an Apache community for Open Climate Workbench and connect to other relevant Apache efforts (Tika, Hadoop, SIS, OODT)
2. Make an initial release.
3. Add new contributors to the project.

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

None at this time.

How has the community developed since the last report?

We're up to 82% (23) of the initial project members having ICLAs on file and accounts. We've got 18% more (5) people still who have yet to submit their ICLA. Chris Douglas and Chris Mattmann have reached out to Chris Jack and Bruce Hewitson and Chris Jack has submitted his ICLA. We are working on processing his account request atm.

How has the project developed since the last report?

- * Primary development has been moved to Apache from JPL now.
- * Previously planned development has been migrated to Apache as well via JIRA issues
- * Cameron Goodale focused on refactoring the core code and breaking it down into smaller functions and objects
- * Mike Joyce and Shakeh Khudikyan have impressive GUI code on its way to the source.
- * Suresh Marru working on refactoring source code to use ASF package names
- * Huikyo Lee now contributing to the discussion on list.

Please check this [X] when you have filled in the report for Open Climate Workbench.

Signed-off-by:

Chris Mattmann: [X](openclimateworkbench)
Suresh Marru: [X](openclimateworkbench)
Chris Douglas: [X](openclimateworkbench)
Nick Kew: [](openclimateworkbench)

Provisionr

Provisionr provides a service to manage pools of virtual machines on multiple clouds.

Provisionr has been incubating since 2013-03-07.

Three most important issues to address in the move towards graduation:

1. release 0.4.0-incubating compliant with the ASF policies
2. develop a community
3. improve support for golden images & cloudstack

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

None at this time.

How has the community developed since the last report?

The community is still bootstrapping the project.

How has the project developed since the last report?

All the project infrastructure has been set up now. Meanwhile development activity has been fairly light so far.

Please check this when you have filled in the report for Provisionr.

Signed-off-by:

Roman Shaposhnik: [](provisionr)

Tom White: (provisionr)

Mohammad Nour El-Din: (provisionr)

Shepherd notes:

Tajo

Tajo is a distributed data warehouse system for Hadoop.

Tajo has been incubating since 2013-03-07.

Three most important issues to address in the move towards graduation:

1. Make an initial Tajo release
2. Grow the Apache Tajo community
3. Foster more committers

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

No

How has the community developed since the last report?

- * A Seoul Tajo meetup was held April 4, 2013 and Jihoon, Sangwook, Jinho, Jaehwa, people who are interested in Tajo, were present. Hyunsik wrote up a summary detailing the meeting, which included an overview of Tajo and briefly mentioning that the roadmap is on the Tajo wiki. Coordinated with ComDev.
- * The number of subscribers for tajo-dev grows up to 40.
- * The Tajo official logo is being voted in <https://issues.apache.org/jira/browse/TAJO-55>.
- * 4 GSoC projects have been proposed.

How has the project developed since the last report?

- * Many Jira issues have been created.
- * 10 Jira issues, including bug fixes and improvements, are resolved.

Please check this [] when you have filled in the report for Tajo.

Signed-off-by:

Chris Mattmann: (tajo)

Owen O'Malley: [](tajo)

Alex Karasulu: (tajo)

Shepherd notes:

Tez

Tez is an effort to develop a generic application framework which can be used to process arbitrarily complex data-processing tasks and also a re-usable set of data-processing primitives which can be used by other projects.

Tez has been incubating since 2013-02-24.

Three most important issues to address in the move towards graduation:

1. Develop collaborations with other Apache projects, including Hadoop, YARN
2. Make an initial Tez release.
3. Grow the Apache Tez community.

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?

None at this time.

How has the community developed since the last report?

The community is still bootstrapping.

How has the project developed since the last report?

A few jiras have been created and some patches have been filed.

Mailing lists for dev, private, and commits are up and running.

There have been some JIRAs, Git is up and running and a couple of larger code contributions have been made.

Chris raised this as a potential issue. In the meanwhile, moderate-low development is occurring.

Please check this when you have filled in the report for Tez.

Signed-off-by:

Alan Gates: (tez)

Arun Murthy: (tez)

Chris Douglas: (tez)

Chris Mattmann: (tez)

Jakob Homan: (tez)

Owen O'Malley: (tez)

Shepherd notes: