

# July2013

⚠ FINAL ⚠

This report has been submitted to the Board. Do not edit further.

Incubator PMC report for July 2013

The Apache Incubator is the entry path into the ASF for projects and codebases wishing to become part of the Foundation's efforts.

The Incubator expanded to 37 podlings this month.

## \* Community

New IPMC members:

(None)

People who left the IPMC:

Davanum Srinivas

New IPMC Chair:

Marvin Humphrey, replacing Benson Margulies

## \* New Podlings

Four new podlings entered the Incubator:

MetaModel  
Olingo  
Spark  
Stratos

## \* Graduations

The board has motions for the following:

Apache Mesos  
Apache JSPWiki

## \* Releases

The following releases were made since the last Incubator report:

Jun 20 Apache Mesos 0.12.0-incubating  
Jun 20 Apache jclouds 1.6.1-incubating  
Jun 22 Apache ODF Toolkit 0.6-incubating  
Jul 02 Apache Curator 2.1.0-incubating

The availability of IPMC votes for podling releases remains a perennial problem and source of discontent. This month, it took between 4 and 20 days from the start of the RC vote on the podling dev list for the third IPMC +1 vote to be cast.

Release	RC VOTE start	Third IMPC +1	Days
Apache Mesos 0.12.0-incubating	Jun 11	Jun 17	6
Apache jclouds 1.6.1-incubating	Jun 13	Jun 17	4
Apache ODF Toolkit 0.6-incubating	May 27	Jun 17	20
Apache Curator 2.1.0-incubating	Jun 20	Jun 28	8

## \* IP Clearance

Cloudant's donation of BigCouch passed by lazy consensus.

\* Miscellaneous

A schism in the IPMC membership was revealed with regards to whether the text of a proposal may be modified while a vote is underway. It was resolved that proposals may not be modified, but opinion remains sharply divided.

Two new draft documentation pages were well received:

- o "What to Expect", an introduction for incoming podling contributors.
- o "Mailing List Options", guidance for requesting mailing lists.

The following topics were discussed on general@incubator but have not yet resulted in concrete action:

- o Provisional top-level projects ("pTLP" for short), a less radical offshoot of the "Incubator Deconstruction" proposal.
- o The language of the Incubator's IP Clearance template (\*not\* the software grant form, only the checklist) is out of date.
- o A "Podling Bill of Rights" was proposed. It has not been formally adopted, but discussion influenced the subsequently spawned "What to Expect" document.
- o A formal ombud position for the Incubator was proposed, and its composition and potential responsibilities were discussed.
- o Exit interviews for graduated podlings were proposed. Debate centered on whether responses should/could/must be anonymous.

An audit by the Chair revealed that autogenerated report reminders sent on July 1 failed to make it to the mailing lists of three podlings (Celix, Tajo, VXQuery), presumably because the emails did not make it through moderation. Manual reminders were sent on July 3 (after subscribing temporarily to the dev lists in question to sidestep moderation).

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

Still getting started at the Incubator (3 podlings)

MetaModel, Spark, Stratos

Not yet ready to graduate (12 podlings)

No release: DeviceMap, Falcon, Knox, MRQL, Open Climate Workbench,  
Ripple, Tajo, VXQuery  
Community growth: Celix, Marmotta, ODF Toolkit, Provisionr

Ready to graduate (5 podlings)

JSPWiki, Mesos, Chukwa, jclouds, Helix

Did not report (2 podlings)

Tashi, Tez

-----  
Table of Contents

Celix  
Chukwa  
DeviceMap  
Falcon  
Helix  
jclouds  
JSPWiki  
Knox  
Marmotta  
MetaModel  
MRQL  
ODF Toolkit  
Open Climate Workbench  
Provisionr  
Ripple  
Spark

Stratos  
Tajo  
Tashi  
Tez  
VXQuery

-----  
  
-----  
Celix

Celix is an OSGi like implementation in C with a distinct focus on interoperability between Java and C.

Celix has been incubating since 2010-11-02.

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

1. Grow the community
2. Find more committers
- 3.

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?

Activity on the mailing list is still a bit slow, but there is still progress. But more importantly, the Native-OSGi effort has taken a few big steps. The most important one is the publication of an RFP at the OSGi Alliance[1]. The purpose of this RFP is to standardise OSGi for native (C and C++) languages. After the RFP is finalised work will be started on a RFC, this will be done in conjunction with writing a reference implementation. For this purpose Apache Celix will be used. This also means that C++ will be introduced in the Celix project. Most likely large parts of C++ code will be donated by 2 participants of the Native-OSGi project. This gives us a great opportunity to add extra committers and increase the diversity of Celix.

How has the project developed since the last report?

As mentioned above, to be able to write a reference implementation for Native-OSGi the scope of Celix will be extended to also include C++. Details of this extension still need to be worked out. But for continuity the current (C) code base will (for now) remain the same. So any work related to Native-OSGi will (for now) be done independently from the current implementation.

Also work is underway to make a second release. Whereas the first release focused primarily on the releasing itself, this second release will focus on the actual implementation. Important points are:

- \* API consistency
- \* Win32 support
- \* APR usage

More details can be found in [2]

A (Mini) Hackathon for Apache Celix was organised at June the 30th. There where 4 people present on the hackathon (2 committers) and focus for the hackathon was the remote services implementation. The hackathon was received well and we like to do this more often to encourage community growth of Apache Celix.

Apache Celix is now also involved the Google Summer of Code for 2013. Erik Jansman is doing a project on the Event Admin for Apache Celix and Pepijn Noltes is mentoring him. The idea is to discuss the result on the mailing list. This is a new experience for Apache Celix and we hope that helps the

community and also provide technical progress.

Date of last release:

16-12-2012

[1]: [https://www.osgi.org/bugzilla/show\\_bug.cgi?id=165](https://www.osgi.org/bugzilla/show_bug.cgi?id=165)

[2]: <http://incubator.markmail.org/thread/fza43iypjutcw2rm>

Signed-off-by:

[x](celix) Marcel Offermans

[ ](celix) Karl Pauls

Shepherd notes:

-----  
Chukwa

Chukwa is a log collection and analysis framework for Apache Hadoop clusters.

Chukwa has been incubating since 2010-07-14.

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

1. Generating an active community
2. Keep code up-to-date with Hadoop trend
3. Make another release

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

Chukwa code needs to be maintained to keep up with new versions of Hadoop.

How has the community developed since the last report?

Committers are communicating more actively than previous quarter.

How has the project developed since the last report?

More contributors have open JIRA to request features and experiment with code base. Website is now svnpubsub enabled. Team member list is updated with new committers. Javadocs are regenerated to remove frame injection vulnerability.

Date of last release:

January 2012

Signed-off-by:

[X](chukwa) Alan Cabrera

[ ](chukwa) Anthony Elder

[ ](chukwa) Bernd Fondermann

[ ](chukwa) Jukka Zitting

Shepherd notes:

ant: I'm a mentor not sheperd but i'll add some comments here. I think we should just graduate Chukwa. They have met all documented Incubator requirements, the only reason they're not graduated is that there isn't loads of development activity going on. Other low activity podlings have graduated before, Wink is a good example, or Steve which didn't even incubate. Chukwa has been incubating for 3 years so aren't just some fly-by-night thing, the "How has the project developed since the last report?" section of this report shows there is work going on, and that they fixed the javadoc frame injection security vulnerability shows they do jump to it when issues that need to be addressed arise. We're not adding any value to the podling by keeping it here in the Incubator, we should graduate them, perhaps including the mentors on the PMC if that helps alleviate any worries.

-----

## DeviceMap

Apache DeviceMap is a data repository containing device information, images and other relevant information for all sorts of mobile devices, e.g. smartphones and tablets. While the focus is initially on that data, APIs will also be created to use and manage it.

DeviceMap has been incubating since 2012-01-03.

There are no issues that the Incubator PMC (IPMC) or the ASF Board need to be aware of.

Our last report indicated a need to reboot the podling, this has been successfully done.

Three new committers and PPMC members have been elected since April, and have been quite active in expanding and improving the podling's code base.

New Java and C# modules have been added and tests are ongoing to validate and compare the various implementations.

The community stays small but is now regularly active, making a release or two in the next period should allow DeviceMap to graduate.

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

1. Make at least one release
2. Clarify how/where DeviceMap gets its device data. Not strictly required to graduate but good for the viability of the project, various options are being evaluated.
3. Graduate - apart from having made a release the podling seems ready for that.

Date of last release:

none yet.

Signed-off-by:

[X](devicemap) Bertrand Delacretaz  
[ ](devicemap) Kevan Miller  
[ ](devicemap) Andrew Savory

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. Add new and diverse committers
2. Build and grow community
3. 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?

\* We had a meetup of Falcon Users along side Hadoop Summit @ San Jose (Jun

25th 2013)

\* More users have joined the falcon users group and mailing lists

How has the project developed since the last report?

9 new JIRAs were created since the last report and 5 of them have been resolved. We are in the process of preparing for our first release since incubation.

Signed-off-by:

[ ](falcon) Arun Murthy  
[X](falcon) Chris Douglas  
[ ](falcon) Owen O'Malley  
[ ](falcon) Devaraj Das  
[ ](falcon) Alan Gates

Shepherd notes:

-----

Helix

Apache Helix is a generic cluster management framework used to build distributed systems and provides automatic partition management, fault tolerance and elasticity.

Helix has been incubating since 2012-10-14.

We feel we are good to graduate, we have had two smooth releases and are working on next release 0.6.2 soon.

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?

-- Added 2 committers Shirshanka Das( LinkedIn) and Vinayak Borkar (UC Irvine)  
-- Good activity on mailing list.  
-- Increase in adoption of Helix outside of LinkedIn

How has the project developed since the last report?

-- Released 0.6.1-incubating with lot of new features.  
-- Issues: 54 created, 25 resolved  
-- Next release 0.6.2 will focus of api refactoring without adding major features.

Date of last release:

05/30/2013 (0.6.1-incubating)

Signed-off-by:

[x](helix) Patrick Hunt  
[x](helix) Olivier Lamy  
[x](helix) Mahadev Konar  
[ ](helix) Owen O'Malley

Shepherd notes:

Suresh Marru: The podling is healthy and I echo it seems ready to graduate. They have cut few releases, are engaging well with contributors and voted in two new committers. Its a small group with activity on the low end but the diversity is decent enough and podling seems to be ready for self-governance. If the project decides to graduate, it might be advisable for one or two mentors continue through for a little while.

-----

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:

1. Garnering additional committers and growing our community
2. Now that we've been through a first successful release, ensuring that subsequent releases progress smoothly and can be handled by multiple members of the community
3. Agreeing upon a new release schedule and roadmap input gathering strategy as part of our first major Apache release

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

Not at the current time. Feedback on the next release will indicate whether the work we've done around license headers is sufficient.

How has the community developed since the last report?

No changes in committer list. Still plenty of activity from existing committers as well as submissions from "new faces" across many of the different jclouds components.

How has the project developed since the last report?

- \* A very important milestone: our first Apache release!
- \* First meetup as an incubator project coming up:  
<http://www.meetup.com/jclouds/events/126378842/>

Date of last release:

Jun 19, 2013 (1.6.1-incubating)

Signed-off-by:

```
[ ](jclouds) Brian McCallister
[X](jclouds) Tom White
[ ](jclouds) Henning Schmiedehausen
[X](jclouds) David Nalley
[ ](jclouds) Jean-Baptiste Onofré
[ ](jclouds) Mohammad Nour El-Din
[ ](jclouds) Olivier Lamy
[ ](jclouds) Tomaz Muraus
[X](jclouds) Suresh Maru
[ ](jclouds) Carlos Sanchez
```

Shepherd notes:

adc: I think this podling is good to go as far as activity and community diversity goes. It's hard to discover what their real status since they've neglected to keep <http://incubator.apache.org/projects/jclouds.html> up to date, but it seems to me that none of the issues listed above are a barrier to their graduation.

It's troubling to see "THE JCLOUDS LOGO IS TRADEMARK OF JCLOUDS, INC." on the Apache JClouds incubator home page. Either the trademark should be assigned to the ASF or it should be removed altogether.

-----

JSPWiki

Java-based wiki engine

JSPWiki has been incubating since 2007-09-17.

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

1. IPMC vote (?, currently running)
- 2.
- 3.

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

- \* We currently have open some JIRAs with Infra. Specifically, having INFRA-5588 (Ubuntu VM for JSPWiki) solved would be a great +1 in terms of community building, and would be a really nice way of entering TLP.

How has the community developed since the last report?

- \* Community graduation vote passed with 16 +1 (<http://s.apache.org/uM>)
- \* There was a call for completing i18n files, which has been fastly addressed by community: Chinese, Portuguese, French, Italian, Brazilian and Russian language files have been completed completed by community
- \* Mail activity has tripled previous quarter at jspwiki-dev, 2.9.1-incubating release and graduation being responsible of part of the increase of this activity. jspwiki-user traffic has decreased a bit, though.

How has the project developed since the last report?

- \* Apache JSPWiki 2.9.1-incubating released on May, 15th
- \* 8 JIRAs resolved in this period
- \* Switch to Maven as build tool almost complete
- \* Upgraded some libraries

Date of last release:

15/05/2013

Signed-off-by:

```
[ ](jspwiki) Dave Johnson
[ ](jspwiki) Craig Russell
[ ](jspwiki) Henning Schmiedehausen
[ ](jspwiki) Sam Ruby
```

Shepherd notes:

wave: JSPWiki is having a graduation vote which just moved to general@. I am +1. Well done! No report as of 7/7.

-----  
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. Voting initiated to invite a new committer to the project.
2. Increased interest in participation resulting from Hadoop Summit exposure.

How has the project developed since the last report?



1. Discussing a 0.3.0 release in the next few weeks.
2. Resolved 26 of 82 total issues currently in JIRA.

Date of last release:

Signed-off-by:

[X](knox) Owen O'Malley  
[X](knox) Chris Douglas  
[ ](knox) Mahadev Konar  
[ ](knox) Alan Gates  
[ ](knox) Devaraj Das  
[X](knox) Chris Mattmann  
[X](knox) Tom White

Shepherd notes:

marvin: The dev list is active and at least one Mentor is engaged.

-----  
Marmotta

An Open Platform for Linked Data.

Marmotta has been incubating since 2012-12-03.

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

1. Ingest code and clear IP.
2. A release
3. Build dev and PMC.

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

No such issues.

How has the community developed since the last report?

- \* dev@marmotta.i.a.o continues with high traffic but slightly decreasing:  
Apr. 265, May 161, Jun. 137. We expect the numbers to raise again with the preparation of the next release.
- \* Two new persons have been elected as new committer and PPMC members:  
Raffaele Palmieri (Eustema Spa) on May 21st, 2013, and Peter Ansell (University of Queensland) on Jun 24th, 2013.

How has the project developed since the last report?

- \* The first release (3.0.0-incubating) has been published on April 24th, 2013. Since that the project is actively working of different improvements, bug fixes and new features, registering 266 issues at jira (60% resolved).
- \* The next release (3.1.0-incubation) is planned for the upcoming weeks  
([http://mail-archives.apache.org/mod\\_mbox/incubator-marmotta-dev/201307.mbox/%3C51D40104.3060203%40apache.org%3E](http://mail-archives.apache.org/mod_mbox/incubator-marmotta-dev/201307.mbox/%3C51D40104.3060203%40apache.org%3E)).
- \* The first two most important issues the project targets to move towards graduation have been already achieved; and the third one, dev community, is being addressed too. Therefore the project has already started to discuss about such option  
([http://mail-archives.apache.org/mod\\_mbox/incubator-marmotta-dev/201305.mbox/%3C51A35FFA.2070707%40apache.org%3E](http://mail-archives.apache.org/mod_mbox/incubator-marmotta-dev/201305.mbox/%3C51A35FFA.2070707%40apache.org%3E)).

Date of last release:

April 24th, 2013.

Signed-off-by:

[ ](marmotta) Fabian Christ  
[x](marmotta) Nandana Mihindukulasooriya  
[x](marmotta) Andy Seaborne

Shepherd notes:

-----

MetaModel

MetaModel is a data access framework, providing a common interface for exploration and querying of different types of datastores.

MetaModel has been incubating since 2013-06-12.

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

1. Complete INFRA setup to move the code to ASF git repository (pending INFRA action): <https://issues.apache.org/jira/browse/INFRA-6404>
2. Redirect MetaModel website to Apache one once available.
3. Bring more diverse committers and contributors to the project.

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?

We have moved the community discussions to [dev@metamodel.incubator.apache.org](mailto:dev@metamodel.incubator.apache.org) to have more open and engaging discussions.

How has the project developed since the last report?

Since the Incubator accepted MetaModel, there's obviously been a lot of attention around the project.

The tech lead for the project, Kasper, had posted some mentionings on [TheServerSide.com](http://TheServerSide.com) and is putting the final touches on an article for [InfoQ.com](http://InfoQ.com) as well (expected published later this week).

The base code has cut another release  
(<http://eobjects.org/trac/query?milestone=MetaModel+3.4.4>)

We are still waiting for INFRA work  
(<https://issues.apache.org/jira/browse/INFRA-6404>) to move the code to ASF git repo.

Date of last release:

None.

Signed-off-by:

[x](metamodel) Henry Saputra  
[x](metamodel) Arvind Prabhakar  
[x](metamodel) Matt Franklin  
[x](metamodel) Noah Slater

Shepherd notes:

-----

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. Complete the first release
  - a) Ensure proper transfer of code
  - b) Verify distribution rights
2. Use maven as our project management tool

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?

No new PPMC members or committers added since last report, We are still working on our first release. We believe that, as soon as we make the first release available, many users will start using the system and will become new contributors.

How has the project developed since the last report?

We have sent a request to trademarks@apache.org to review and verify our searches that establish that "Apache MRQL" is a suitable name. More specifically, our US and EU trademark searches, and a Google search, show that the project name MRQL does not conflict with a previously registered trademark. We are now in the process of completing the MRQL codebase to make it ready for the first release. There are two component missing that delay this release: switching to maven as our project management tool and verifying copyright information.

Date of last release:

none yet

Signed-off-by:

[X](mrql) Alan Cabrera  
[X](mrql) Alex Karasulu  
[ ](mrql) Mohammad Nour El-Din

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. Hold an informal graduation readiness vote

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?

An uptick in activity related to our most-recent release, but no new developers.

Discussion of whether we should aim for TLP status or merge into an existing TLP with similar alignment, such as POI or OpenOffice.

How has the project developed since the last report?

Successful release of version 0.6

Date of last release:

2013-06-22

Signed-off-by:

```
[ ](odftoolkit) Sam Ruby
[ ](odftoolkit) Nick Burch
[ ](odftoolkit) Yegor Kozlov
```

Shepherd notes:

adc: The report seems pretty accurate.

-----  
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?

Apache Open Climate Workbench elected its first new committer and PPMC member, Alex Goodman, in the past month.

How has the project developed since the last report?

- \* Cameron Goodale is working to push the 0.1-incubating release out the door (currently on RC #4, with #5 coming soon)
- \* Tons of development going on (already up to nearly 200 JIRA issues; ReviewBoard actively being used)
- \* The UI is receiving the most work in particular by Michael Joyce and Shakeh Khudikyan.
- \* Kyo Lee and Alex Goodman and Kim Whitehall and Jinwon Kim are working on the metrics module.
- \* Documentation is now appearing on the wiki including a guide produced by Alex Goodman for Easy RCMET, our installer.
- \* Maziyar Boustani has several patch reviews going on to suggest a new API for post 0.1-incubating development for RCMET.

Date of last release:

Signed-off-by:

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

Shepherd notes:

adc: The missing report is in line with this project's weakness, filling out paperwork. ;) The podling status page <http://incubator.apache.org/projects/climate.html> needs to be filled out.

With that said, the project seems pretty active with more than one developer doing the heavy lifting.

Again, it would be good to fill out the status page so we can see where they're at since they may be good candidates for graduation soon.

marvin: (Alan's notes were written up before OCW filed a late report.)

-----  
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. Gather feedback on the first release and plan a new one
2. Build a community
3. Document all the important bits

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?

Some new folks on the mailing list (20 subscribers)

How has the project developed since the last report?

Our first release as an Apache Incubator project is out (0.4.0-incubating)

Date of last release:

July 2, 2013

Signed-off-by:

[ ](provisionr) Roman Shaposhnik  
[X](provisionr) Tom White  
[ ](provisionr) Mohammad Nour El-Din

Shepherd notes:

-----  
Ripple

Ripple is a browser based mobile phone emulator designed to aid in the development of HTML5 based mobile applications. Ripple is a cross platform and cross runtime testing/debugging tool. It currently supports such runtimes as Cordova, WebWorks aand the Mobile Web.

Ripple has been incubating since 2012-10-16.

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

1. Check if some source modules are not compatible with ASF guidelines and need to be extracted to f. e. GitHub
2. Check license of images and - if necessary - replace them
3. Prepare first release

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?

More issue reports, feature suggestions and general questions submitted by users.

More code contributions from individual users (including instances where the user who submitted an issue went and fixed it themselves).

How has the project developed since the last report?

More activity via non-committer contributions (alongside part time committer contributions). Mainly: bug fixes.

Signed-off-by:

[ ](ripple) Jukka Zitting  
[X](ripple) Christian Grobmeier  
[ ](ripple) Andrew Savory

Shepherd notes:

-----

Spark

Spark is an open source system for fast and flexible large-scale data analysis. Spark provides a general purpose runtime that supports low-latency execution in several forms.

Spark has been incubating since 2013-06-19.

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

1. Finish bringing up infrastructure on Apache (JIRA, "user" mailing list, SVN repo for website)
2. Migrate mailing lists and development to Apache
3. Make a Spark 0.8 under the Apache Incubator

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

While most of our infrastructure is now up, it has taken a while to get a JIRA, a SVN repo for our website (so we can use the CMS), and a user@spark.incubator.apache.org mailing list (so we can move our existing user list, which is large).

How has the community developed since the last report?

We only entered the Apache Incubator at the end of June, but in the existing developer community keeps expanding and we are seeing many new features from new contributors.

How has the project developed since the last report?

In terms of the Apache incubation process, we filed our IP papers and got a decent part of the infrastructure set up (Git, dev list, wiki, Jenkins group).

Date of last release:

None

Signed-off-by:

[X](spark) Chris Mattmann  
[ ](spark) Paul Ramirez  
[ ](spark) Andrew Hart  
[ ](spark) Thomas Dudziak  
[X](spark) Suresh Marru  
[x](spark) Henry Saputra  
[ ](spark) Roman Shaposhnik

Shepherd notes:

-----

Stratos

Stratos will be a polyglot PaaS framework, providing developers a cloud-based environment for developing, testing, and running scalable applications, and IT providers high utilization rates, automated resource management, and platform-wide insight including monitoring and billing.

Stratos has been incubating since 2013-06-20.

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

1. Finish infrastructure setup for the project
2. Perform IP audit of the codebase
3. Perform a release

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 have moved the code into the ASF repo.

Plan on educating the community on what the current design and structure of the project's code

The milestone plan proposed for next immediate release so that the community can start engaging

How has the project developed since the last report?

This is our first report as a podling

Date of last release:

No release yet

Signed-off-by:

[ ](stratos) Afkham Azeez  
[x](stratos) Anthony Elder  
[X](stratos) Chip Childers  
[ ](stratos) Marlon Pierce  
[ ](stratos) Mohammad Nour  
[x](stratos) Noah Slater  
[x](stratos) Suresh Maru

Shepherd notes:

wave: The project is just starting up and still getting infrastructure in place. Looks to be excellent energy with the usual impatience with infrastructure. Should the IPMC and Infra collaborate on providing a timeline to set podling expectations?

-----

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. Ensure at least one Tajo release
2. Attract more contributors and committers
3. Grow Tajo 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 are mentoring two GSoC students on Tajo (Hybrid Hash Join and Outer Join issues)
- \* We decided the official Tajo logo.
- \* Reviewboard was set up.

How has the project developed since the last report?

- \* The package name is changed to org.apache.tajo.
- \* We have improved SQL parser to support more SQL standard.

Date of last release:

Signed-off-by:

```
[ ](tajo) Chris Mattmann
[ ](tajo) Owen O'Malley
[ ](tajo) Alex Karasulu
```

Shepherd notes:

marvin: Dev list is active and at least one Mentor is participating.

-----  
Tashi

An infrastructure for cloud computing on big data.

Tashi has been incubating since 2008-09-04.

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?

Date of last release:

Signed-off-by:

```
[ ](tashi) Matthieu Riou
[ ](tashi) Craig Russell
```

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.
- 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?

Date of last release:

Signed-off-by:

```
[ ](tez) Alan Gates
[ ](tez) Arun Murthy
[ ](tez) Chris Douglas
[ ](tez) Chris Mattmann
[ ](tez) Jakob Homan
[ ](tez) Owen O'Malley
```

Shepherd notes:

marvin: Active and Mentors are engaged. The dev@ list is mostly Jenkins spam and all the action is on the issues@ and commits@ lists, which may be surprising to people who expect that "if it didn't happen on the dev list it didn't happen".

-----  
VXQuery

A standards compliant parallel XML Query processor.

VXQuery has been incubating since 2009-07-06.

The most important issue to address in the move towards graduation is still the first release. Two release candidates have been created and rejected so far. The third release candidate is not ready yet.

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

It would be great to have more active mentors.

How has the community developed since the last report?

Steven Jacobs has started to contribute to the project as part of his GSoC project of integrating Lucene indices into VXQuery.

How has the project developed since the last report?

Work on a benchmark to evaluate the performance and applicability of a parallel XQuery engine has been started. This has triggered further progress on the rewriting of queries to increase the level of parallelism for some classes of queries.

Signed-off-by:

```
[ ](vxquery) Sanjiva Weerawarana
[ ](vxquery) Radu Preotiu-Pietro
[ ](vxquery) Jochen Wiedmann
```

Shepherd notes: