

January2015

/!\ *FINAL* (!)

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

Incubator PMC report for January 2015

Timeline

Wed January 07	Podling reports due by end of day
Sun January 11	Shepherd reviews due by end of day
Sun January 11	Summary due by end of day
Tue January 13	Mentor signoff due by end of day
Wed January 14	Report submitted to Board
Wed January 21	Board meeting

Shepherd Assignments

Alan D. Cabrera	Ignite
Alan D. Cabrera	Lens
Andrei Savu	Kylin
Andrei Savu	Usergrid
Dave Fisher	Kalumet
Drew Farris	DataFu
John Ament	HTrace
John Ament	Ranger
Justin Mclean	Taverna
Konstantin Boudnik	MRQL
Konstantin Boudnik	Parquet
Matthew Franklin	Calcite
Matthew Franklin	Samza
P. Taylor Goetz	Aurora
P. Taylor Goetz	NPanday
Raphael Bircher	Corinthia
Raphael Bircher	SAMOA
Ross Gardler	Ripple
Suresh Marru	Brooklyn
Suresh Marru	Tamaya
Timothy Chen	NiFi
Timothy Chen	ODF Toolkit

Report content

Incubator PMC report for January 2015

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

There are currently 36 podlings undergoing incubation. One podling joined us this month (Corinthia). One member joined and two members left the IPMC.

IPMC has recognized the need for tightening up mentorship requirements and overall structure of the incubation process. Active discussions on how to this in the best possible way are on going and the recommendation is expected to be available in a few weeks.

* Community

New IPMC members:

Hyunsik Choi

People who left the IPMC:

Sean Owen

Marvin Humphrey

Mentors who didn't sign off on the reports:

Alan Gates
Alex Karasulu
Andrei Savu
Andrew Purtell
Arun Murthy
Ashutosh Chauhan
Benjamin Hindman
David Blevins
Devaraj Das
Drew Farris
Enis Soztutar
Gerhard Petracek
Henri Gomez
Lewis John Mcgibbney
Luciano Resende
Marcel Offermans
Matt Hogstrom
Nick Burch
Olivier Lamy
Sam Ruby
Sergio Fernandez
Suresh Maru
Suresh Srinivas
Todd Lipcon
Tom White
Yegor Kozlov

* New Podlings

Corinthia
Zeppelin

* Graduations

The board has motions for the following:

Samza

* Releases

The following releases were made since the last Incubator report:

Dec 05 2014 Apache Falcon 0.6-incubating

Dec 08 2014 Apache Samza 0.8.0-incubating
Dec 22 2014 Apache Brooklyn 0.7.0-M2-incubating

* IP Clearance

* Corinthia initial source grant

* Legal / Trademarks

* Infrastructure

* Miscellaneous

* NPanday community seems to be in agreement that retirement is the best option at this point. The only outstanding issue before formally recommending graduation VOTE is to decide whether there's enough cycles available for one last release before retirement.

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

* Still getting started at the Incubator

SAMOA
Corinthia
Kylin
NiFi
Taverna (delayed software grant)
Zeppelin

* Not yet ready to graduate

No release:

DataFu
HTrace
Ignite
Kalumet
Lens
Tamaya

Community growth:

Aurora
Brooklyn
Calcite
MRQL
ODF Toolkit
Parquet
Ranger
Usergrid

* Ready to graduate

The Board has motions for the following:

Samza

* Did not report, expected next month

NPanday
Ripple

Table of Contents

Aurora
Brooklyn
Calcite

Corinthia
DataFu
HTrace
Ignite
Kalumet
Kylin
Lens
MRQL
NiFi
NPanday
ODF Toolkit
Parquet
Ranger
SAMOA
Samza
Tamaya
Taverna
Usergrid
Zeppelin

Aurora

Aurora is a service scheduler used to schedule jobs onto Apache Mesos.

Aurora has been incubating since 2013-10-01.

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

1. Expanding the community's diversity and adding new committers.
2. Third Apache release, progress being tracked in ticket AURORA-872
- 3.

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?

Latest Additions:

- * PMC addition:
 - * David McLaughlin, 2014-08-20
- * Contributor additions: No additional contributors had code committed since the last report, however three new contributors currently have code under review that was submitted within the past 6 weeks. At least one of these contributions is a large feature addition, which is promising.

Issue backlog status since last report (Oct 1, 2014):

- * Created: 212
- * Resolved: 177

Mailing list activity since last report:

- * @dev 480 messages
- * @issues 1085 messages
- * @reviews 2235 messages

How has the project developed since the last report?

- * Increased participation from new contributors within the Aurora core and ecosystem, for example a Chef cookbook and Docker support under development.
- * Continued to hold weekly IRC meetings, digests sent to dev@ list
- * Large improvements to project documentation and website.

Date of last release:

* Apache Aurora 0.6.0 (incubating) - 2014-11-21

When were the last committers or PMC members elected?

* David McLaughlin, 2014-08-20

Signed-off-by:

[X](aurora) Jake Farrell
[](aurora) Benjamin Hindman
[X](aurora) Chris Mattmann
[X](aurora) Henry Saputra

Shepherd/Mentor notes:

P. Taylor Goetz (ptgoetz):

Podling looks very healthy. 3 of 4 mentors engaged since last report.

Brooklyn

Brooklyn is a framework for modelling, monitoring, and managing applications through autonomic blueprints.

Brooklyn has been incubating since 2014-05-01.

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

1. Grow the community
2. More diversity of the committers/PPMC (currently biased towards employees of a single organization)
3. Following up our first release with further, regular, releases

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?

Our community continues to grow slowly but surely, and has received interest and contributions from new community members.

We had the opportunity to talk about our project at ApacheCon Europe recently, which has introduced Brooklyn to a wider audience within Apache.

How has the project developed since the last report?

Code continues to be proposed and merged at a high rate. We are satisfied with our technical velocity.

We recently completed our first release as an Apache project, and look forward to following this up with regular releases from now on.

Date of last release:

2014-12-23 (apache-brooklyn-0.7.0-M2-incubating)

When were the last committers or PMC members elected?

No change since 2014-07-02.

Signed-off-by:

[](brooklyn) Matt Hogstrom
[](brooklyn) Alex Karasulu
[X](brooklyn) David Nalley

[](brooklyn) Marcel Offermans
[X](brooklyn) Jean-Baptiste Onofré
[](brooklyn) Olivier Lamy
[X](brooklyn) Chip Childers
[](brooklyn) Andrei Savu
[X](brooklyn) Joe Brockmeier
[X](brooklyn) Jim Jagielski

Shepherd/Mentor notes:

Chip Childers (chipchilders):

The Brooklyn community is doing very well in learning the rules / policies / nuances of working at the ASF. It still needs a bit more time to build up a diverse community IMO, but things are looking good.

David Nalley (ke4qqq):

The Brooklyn folks kicked out their first release during this reporting cycle, which is an important milestone. Work still remains in terms of building a diverse community.

Calcite

Calcite is a highly customizable engine for parsing and planning queries on data in a wide variety of formats. It allows database-like access, and in particular a SQL interface and advanced query optimization, for data not residing in a traditional database. (Renamed from Optiq on 2014-09-30.)

Calcite has been incubating since 2014-05-19.

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

1. Continue to build community
2. Add more committers and PPMC members
3. Demonstrate community robustness by rotating project tasks among multiple project members

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?

In December, there were 84 messages on dev@, 45 commits, 50 issues created, 27 issues resolved by 7 contributors.

Julian Hyde gave a talk "SQL on everything, in memory" about Calcite at the Strata + Hadoop World conference in October, and held an "office hour" with interested community members. He also spoke at a Hive meet-up in New York and a Kylin meet-up in San Jose about how those projects integrate with Calcite.

Collaborations are under way with other Apache projects: with Drill to make implementation convention more neutral; with Hive for join algorithm selection, and leveraging sorted data; and with Phoenix to create a Calcite adapter for Phoenix.

How has the project developed since the last report?

We have completed the rename from Optiq to Calcite. We have also re-organized the code into the org.apache.calcite namespace. We are planning to make a 1.0 release in January with the re-organized code.

Date of last release:

2014-11-05 (0.9.2)

When were the last committers or PMC members elected?

PPMC member elected in December (but has not yet accepted).

Signed-off-by:

[X](calcite) Ted Dunning
[X](calcite) Alan Gates
[X](calcite) Steven Noels

Corinthia

Corinthia is a toolkit/application for converting between and editing common office file formats, with an initial focus on word processing. It is designed to cater for multiple classes of platforms - desktop, web, and mobile - and relies heavily on web technologies such as HTML, CSS, and JavaScript for representing and manipulating documents. The toolkit is small, portable, and flexible, with minimal dependencies. The target audience is developers wishing to include office viewing, conversion, and editing functionality into their applications.

Corinthia has been incubating since 2014-12-08.

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

1. complete the move to ASF infrastructure
2. Discuss and decide on a medium-term roadmap
3. Grow the 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?

This is our first podling report.

We have spend a lot of mail on getting acquainted and finding the strength of each other.

Now discussion are flowing freely. We have consensus on PPMC == Committer as long as we are in incubator.

How has the project developed since the last report?

This is our first podling report.

We are at work making a stable kernel so we can more easily add more developers.

Date of last release:

There have been no releases.

When were the last committers or PMC members elected?

Andrea Pescetti 2014-12-31 added as committer/PPMC

Signed-off-by:

[X](corinthia) Daniel Gruno
[X](corinthia) Jan Iversen

DataFu

DataFu provides a collection of Hadoop MapReduce jobs and functions in higher level languages based on it to perform data analysis. It provides functions

for common statistics tasks (e.g. quantiles, sampling), PageRank, stream sessionization, and set and bag operations. DataFu also provides Hadoop jobs for incremental data processing in MapReduce.

DataFu has been incubating since 2014-01-05.

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

1. Grow user and contributor base.
2. Make first release.
3. Increase activity for initial committers.

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

1. Has not yet made a release, but is in process of preparing first.
2. Need to dramatically grow the contributor base.

How has the community developed since the last report?

New committer and PMC member. Several JIRAs filed by new users.

How has the project developed since the last report?

1. 16 issues created, several from new contributors.
2. 8 issues closed.
3. Reasonable amount of mailing list traffic.

Date of last release:

None yet. Currently preparing release: DATAFU-53.

When were the last committers or PMC members elected?

Nov 2014, Russell Journey, both committer and PPMC.

Signed-off-by:

[](datafu) Ashutosh Chauhan
[X](datafu) Roman Shaposhnik
[X](datafu) Ted Dunning

HTrace

HTrace is a tracing framework intended for use with distributed systems written in java.

HTrace has been incubating since 2014-11.

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

1. Make our first release from incubator
2. Move dependent projects over to the incubator version
3. Continue to grow user and contributor base; i.e. grow the 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?

- * 141 messages on dev@htrace
- * 42 issues resolved and 11 open

How has the project developed since the last report?

- * This is our first report.
- * We are fully moved over to Apache Infrastructure now (thanks in particular to Jake Farrell for help here).

- * We have our first htrace website published.
- * Development of an easy to deploy htraced trace sink and visualization is moving along nicely.
- * An htrace receiver for Apache Flume was contributed.
- * We have put up 5 release candidates for our 1st podling release with a 6th on its way.

Date of last release:

None as yet.

When were the last committers or PMC members elected?

None since initial setup.

Signed-off-by:

```
[X](htrace) Jake Farrell
[ ](htrace) Todd Lipcon
[X](htrace) Lewis John Mcgibbney
[ ](htrace) Andrew Purtell
[X](htrace) Billie Rinaldi
[X](htrace) Michael Stack
```

Shepherd/Mentor notes:

John D. Ament (johndament):

The podling is working through its first release slowly. First release is always the hardest. Good activity for a new podling on the lists, still waiting on the board report though.

Ignite

A unified In-Memory Data Fabric providing high-performance, distributed in-memory data management software layer between various data sources and user applications.

Ignite has been incubating since 2014-10-01.

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

1. Finish migration to Apache process.
2. Grow active and healthy community around Apache Ignite product.
3. Get on a stable release schedule and have 3 successful product releases.

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

None at the moment.

How has the community developed since the last report?

Several contributors expressed interest in contributing.

How has the project developed since the last report?

- Started the implementation of JCache (JSR 107) specification.
- Integrated JCache (JSR 107) TCK test suite into project lifecycle.
- Made significant improvements to Apache Ignite (incubating) website.
- Submitted SGA from GridGain to Apache Ignite (incubating) for In-Memory Hadoop Accelerator.

Date of last release:

No releases yet: the code is still in the migration phase

When were the last committers or PMC members elected?

10/01/2014 Initial committers were added.

Signed-off-by:

[x](ignite) Branko ibej
[x](ignite) Konstantin Boudnik
[X](ignite) Henry Saputra
[X](ignite) Roman Shaposhnik
[X](ignite) Michael Stack

Shepherd/Mentor notes:

Alan Cabrera (adc):

This podling is off to a great start. Mentors are extremely active participants with most of the non-Jira chatter being between Dmitriy Setrakyan and the mentors; look to the Jira issue to see the rest of the community "getting it done". The "issues to be addressed" section seems reasonable.

Branko ibej (brane):

I agree with Alan's assessment, but I'd really like to see more activity on the dev@ list; especially more of the initial committers getting involved in discussions there. Whilst Jira notifications are nice, I have a nagging feeling that most of the discussions within the community are happening off-line, which leaves a rather large hole in the community's collective memory (i.e., mail archives). This will have to be addressed before graduation.

Kalumet

Kalumet a complete environment manager and deployer including J2EE environments (application servers, applications, etc), softwares, and resources.

Kalumet has been incubating since 2011-09-20.

Community Developement:

Apache Kalumet 0.6-incubating version has been released. However, due to some misunderstanding, the release vote has not been completed by 3 IPMC. Especially, some legal files issues have been raised.

We are preparing a 0.6.1-incubating release to fix the legal files and submit to IPMC vote.

We are in the way of promoting the documentation on the website.

Project Development:

We are preparing the 0.6.1-incubating (plan for December, 27) to fix the legal files issues raised on 0.6-incubating release and have a complete IPMC vote.

We are preparing the 0.7-incubating release with the development changes.

A new contributor has been identified and started to work on some issues and refactoring of the webconsole (mockup provided using Vaadin).

Local branches have been created containing:

- new model and REST API
- new webconsole (remove of Echo framework)

These local branches will be merged on the 0.7-incubating branch (master).

Before Graduation:

- The documentation has been updated and aligned with the 0.6-incubating release. The documentation will be promoted on the website and "linked" in announcement e-mails as soon as the 0.6.1-incubating release has been voted.
- The 0.7-incubating release will fix the graduation issue, especially in term of license (removal of Echo2 framework for the UI).

Post Graduation:

- Kalumet scripts. It's the extension of the "software" concept globally to all Kalumet resources. It will allow users to create custom deployment/update plan, with a set of fine-grained steps.

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

None so far.

Date of last release:

2013-11-22

Date of next release:

2014-12-27

When were the last committers or PMC members elected?

None

Signed-off-by:

[X](kalumet) Jim Jagielski
 [](kalumet) Henri Gomez
 [X](kalumet) Jean-Baptiste Onofre
 [](kalumet) Olivier Lamy

 Kylin

Kylin is a distributed and scalable OLAP engine built on Hadoop to support extremely large datasets.

Kylin has been incubating since 2014-11-25.

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

1. Migrate source code and issue tracking from github.com to Apache Git and JIRA
2. Make the first release from apache
3. Make project website available on apache

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

* Nothing

How has the community developed since the last report?

The community is growing. Several new members have started actively contributing to the project. New contributors made a significant contribution in develop new feature, fix bugs, refine web experience. The Kylin meetups at Bay Area and Beijing was well attended. Apache Kylin has been presented at Big Data Technology Conference 2014, Beijing which is one of biggest Big Data conference in China. Interest of people during that event was great and engaged a lots of organizations and individuals, many of them started to evaluate and joined Kylin community.

How has the project developed since the last report?

1. The project setup is going quite well. Initial committers are

onboarding, JIRA, mailing list were successfully done, Web content has been migrated to Apache Git but need tuning. Main source code migration will be done in January 2015.

2. Bug fixing, enhancement are happening on current branch which will be the target v1.0 release.
3. Work toward next gen version (supports inverted-index) is progressing substantially.

Date of last release:

None yet. We are preparing to release 1.0-incubating (MOLAP) within the next 1 month.

When were the last committers or PMC members elected?

NONE

Signed-off-by:

[x](kylin) Owen O'Malley
[x](kylin) Ted Dunning
[X](kylin) Henry Saputra

Lens

Lens is a platform that enables multi-dimensional queries in a unified way over datasets stored in multiple warehouses. Lens integrates Apache Hive with other data warehouses by tiering them together to form logical data cubes.

Lens has been incubating since 2014-10-10.

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

1. Make the first release from apache
2. Make project website available on apache
3. Reach out to people through presentations and blogs

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

We are still waiting for project website creation via
<https://issues.apache.org/jira/browse/INFRA-8549>

How has the community developed since the last report?

- * Logo of the project has been finalized and published.
- * Added one new committer/PPMC - Rajat Khandelwal
- * Had contributor meetup in Bangalore, India
- * Mailing list subscription (as on 2015-jan-03):
 - * Dev list: 29 members
 - * Commits list: 14 members
 - * User list: 21 members

How has the project developed since the last report?

- * Main activities include
 - * Podling setup wrt jenkins and pushing docs
 - * Improving documentation
 - * Stabilizing the code for first stable release
 - * Feature additions - Join chains in OLAP cube
- * Mailing list activity :
 - * Dev via mail-archives.apache.org:
 - * December - 1380 messages
 - * Commits via mail-archives.apache.org:
 - * December - 50 messages
 - * User mailing list activity - None
- * Issues 101 created and 62 resolved (via 30 day summary from LENS Jira)

Date of last release:

NONE

When were the last committers or PMC members elected?

2014-12-24

Signed-off-by:

[X](lens) Christopher Douglas
[X](lens) Jakob Homan
[X](lens) Jean-Baptiste Onofre

Shepherd/Mentor notes:

Alan Cabrera (adc):

Very good report; an exemplar of how reports should be done. The community seems very active and healthy. IMO, only items 1 & 2 are required for graduation.

MRQL

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

MRQL has been incubating since 2013-03-13.

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

1. Increase the number of active committers
2. Increase adoption, expand user community, and increase user list activity

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

The project is not making much progress on the level of increasing the number of active committers and expanding the users community.

This is already being discussed inside the project community and the main active committer asked to give the project more few months as he will try spreading the word more about the project through some public events

How has the community developed since the last report?

None

How has the project developed since the last report?

During the last three month, 12 jira issues were reported, from which 10 were fixed. Most of these issues were related to MRQL query evaluation in Flink mode. Finally, we have set up MRQL on ASF Jenkins, which is a CI server that continuously checks the integrity of our builds and validates our tests.

Date of last release:

2014-06-26

When were the last committers or PMC members elected?

2014-04-17

Signed-off-by:

[X](mrql) Alan Cabrera
[](mrql) Alex Karasulu

[X](mrql) Mohammad Nour El-Din

Shepherd/Mentor notes:

Konstantin Boudnik (cos):

Very light email list traffic compare to prev. month. Mentors are active.

NiFi

NiFi is a dataflow system based on the concepts of flow-based programming.

NiFi has been incubating since 2014-11-24.

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

1. Produce an initial release.
2. Grow our community beyond the initial set of committers.
3. Establish rhythm for releases, review processes, and document how to most effectively bring on new committers.

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

Nothing we wish to raise at this time.

How has the community developed since the last report?

The dev mailing list has picked up 20 new e-mails since the last report and is now at 62 (commits has grown to 25).

At the suggestion of the community we've adopted a review-then-commit strategy which is designed to among other things help with community development.

We've received several patches from members outside the initial committer list including changes which dramatically improved our build process and provide new extensions.

How has the project developed since the last report?

All of our current infrastructure needs are setup and things are working really well. Our focus now is on working towards our first release which is 0.0.1 and presently comprised of 85 JIRA issues the vast majority of which are resolved. We've conducted a thorough review of all source files to ensure proper license headers/notices/disclaimers are present. We've completed a thorough review and modification of the overall LICENSE and NOTICE files in preparation for release. We've conducted the initial podling name search and submitted the ticket for that analysis to be verified and confirmed.

Date of last release:

N/A

When were the last committers or PMC members elected?

We have not grown the official committer list beyond the initial committer list. Though there are a couple of folks who are committing fairly consistently now and a new member to join the dev mailing list that is very promising.

Signed-off-by:

[X](nifi) Billie Rinaldi
[x](nifi) Arvind Prabhakar
[](nifi) Sergio Fernandez
[x](nifi) Benson Margulies
[x](nifi) Brock Noland

[](nifi) Drew Farris
[X](nifi) Andrew Purtell

Shepherd notes (Tim Chen):

Project is still setting up and in the early phase. Mailing list and commits are fairly active.

NPanday

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

NPanday has been incubating since 2010-08-13.

Shepherd/Mentor notes:

P. Taylor Goetz (ptgoetz):

NPanday has missed the last two reports (last report August 2014). One of two mentors are engaged and encouraging the podling to make a final release before retiring (last release was on 2011-05-17). However, there is little PPMC/Committer involvement with only 2 PPMC members showing involvement on the dev@ list in the past year. Also confirmed that a report reminder email was received by the dev@ list for each of the last missed reporting periods as well at the current period.

It's clear that mentors past and present have attempted to help the NPanday project grow to a viable size -- at least enough to approve a release, new Committer/PPMC Member, or even decide to retire -- without success.

The main warning signs appear to be:

- * Inability to reach community consensus (barely 2 active Committer/PPMC Members, making 3 +1s for any vote very difficult)
- * Inability to perform releases
- * Inability to grow the community to a viable size

The issues listed above do not indicate a lack of mentor involvement, but rather a community in decline.

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 in terms of committers
2. Have regular releases
3. Decide if we would like to be a top level project or join an existing project, or move to the attic.

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

No new.

How has the community developed since the last report?

We have made a release but there has been little activity since then.

We still need to attract more members and publicizing the project in the ODF.

Our validator was used on the ODF Plugfest, ahead a patch was provided by

a new contributor.

How has the project developed since the last report?

Update of ODF validator before the ODF Plugfest (London).

There have been some user issues reported and responded to.

ODF Toolkit part of presentation on the Apache Con Europe.

Date of last release:

2014-06-02

When were the last committers or PMC members elected?

2012-10-29

Signed-off-by:

[x](odftoolkit) Rob Weir
[](odftoolkit) Sam Ruby
[](odftoolkit) Nick Burch
[](odftoolkit) Yegor Kozlov

Shepherd/Mentor notes:

Additional mentor comments (from robweir) here:
<http://s.apache.org/bUW>

(Tim Chen)

It seems like the project is active in release and code (towards the low activity side).
However the mailing list traffic and engagement has decreased overtime and the
dev list even with jira notifications is still very low traffic.

Parquet

Parquet is a columnar storage format for Hadoop.

Parquet has been incubating since 2014-05-20 .

Three most important issues

- Expanding the community and adding new committers
- 1st releases toward org.apache Parquet 1.6.0 GA
- Identifying how to ensure timely code reviews by committers

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

- None at this time

Latest Additions:

- * PMC addition: None
- * Contributor addition: Dan Weeks and Alex Levenson

Issue backlog status since last report:

- * Created: 45
- * Resolved: 20

Mailing list activity since last report:

- * dev 310 messages: 90 in Oct, 126 in Nov, and 94 in Dec

How has the project developed since the last report?

- Completed first release, Apache Parquet Format (incubating) 2.2.0
- Established a by-law for adding committers
- Added 2 new committers
- Parquet presentation accepted for Strata San Jose

Date of last release:

- parquet-format released 14 November 2014
- Not yet released: parquet-mr and parquet-cpp

Signed-off-by:

```
[ ](parquet) Todd Lipcon
[X](parquet) Jake Farrell
[X](parquet) Chris Mattmann
[X](parquet) Roman Shaposhnik
[ ](parquet) Tom White
```

Shepherd/Mentor notes:

Mailing lists are active; most mentors are active.

Ranger

The Ranger project is a framework to enable, monitor and manage comprehensive data security across the Hadoop platform.

Apache Ranger has been incubating since 2014-07-24.

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

- * Increase the Ranger contributors/committers across industry verticals
- * Establish a automated git check-in process to validate the check-in patches for quality control (similar to Hadoop QA)
- * Add more documentation to support new 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?

- * Added few more contributors
- * Vote for adding an additional committer is in progress with private@ranger.incubator
- * Issues (JIRAs) created/resolved since last report (after Oct-01-2014 before Dec-30-2014):

Created:	115
Resolved:	100

How has the project developed since the last report?

- * Project Name has been renamed from "Apache Argus" to "Apache Ranger"
- * Three Release candidates were voted in dev community for Ranger-0.4.0 version and RC3 has been voted to be released in IPMC
- * First release of Apache Ranger 0.4.0 on Nov-17,2014
- * Apache Ranger 0.4.0 Release artifacts are published in official dist directory.
- * Ranger Documentation wiki Site (<https://cwiki.apache.org/confluence/display/RANGER>) has been added.
- * Working on release of Apache 0.4.1 to avoid compilation issues related to SNAPSHOTS libraries in HBase and Hive dependencies
- * Working on adding more security functionalities for the next release, Apache Ranger 0.5.0

Date of last release:

Nov-17-2014

When were the last committers or PMC members elected?

none yet

Signed-off-by:

[X](ranger) Alan Gates
[X](ranger) Daniel Gruno
[](ranger) Devaraj Das
[X](ranger) Jakob Homan
[X](ranger) Owen O'Malley

Shepherd/Mentor notes:

John D. Ament (johndament):

Ranger suffered a bit from a renaming issue that came up. Activity on the lists is a little bit lighter since, but I'm sure will bounce back. I'm concerned over inability to access the wiki though, sent a note to the podling to request proper access.

SAMOA

SAMOA provides a collection of distributed streaming algorithms for the most common data mining and machine learning tasks such as classification, clustering, and regression, as well as programming abstractions to develop new algorithms that run on top of distributed stream processing engines (DSPes). It features a pluggable architecture that allows it to run on several DSPes such as Apache Storm, Apache S4, and Apache Samza.

SAMOA has been incubating since 2014-12-15 and is not fully functioning as a project yet.

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

1. Get IP clearance (SGA) from Yahoo
2. Move the current code into ASF's git repository
3. Start working as an Apache project

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 the first report.

How has the project developed since the last report?

This is the first report.

Date of last release:

No incubator release yet.

When were the last committers or PMC members elected?

We just established the initial PPMC.

Signed-off-by:

[](samoa) Alan Gates
[](samoa) Ashutosh Chauhan
[](samoa) Enis Soztutar
[X](samoa) Ted Dunning

Samza

Samza is a stream processing system for running continuous computation on infinite streams of data.

Samza has been incubating since 2013-07-30.

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

None. PPMC and IPMC have voted for graduation. Waiting on Board meeting to pass final resolution.

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?

Continued growth in mailing list, particularly from new users.

How has the project developed since the last report?

Released version 0.8.

Date of last release:

2014-12-09

When were the last committers or PMC members elected?

Chinmay Soman, 2014-08-13, PPMC and committer.

Signed-off-by:

[X](samza) Chris Douglas
[](samza) Arun Murthy
[X](samza) Roman Shaposhnik

Tamaya

Tamaya is a highly flexible configuration solution based on an modular, extensible and injectable key/value based design, which should provide a minimal but extendible modern and functional API leveraging SE, ME and EE environments.

Tamaya has been incubating since 2014-11-14.

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

1. Project web site must be setup (done).
2. Establish a solid code base for a release.
3. Continue to grow the community.

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

There are many discussions about how simple the API and or code should be.

How has the community developed since the last report?

We got a new contributor, Reinhard Sandtner who already shipped a patch and is participating in some discussions.

How has the project developed since the last report?

The project setup is going quite well. The initial code donation is under ongoing review.

Date of last release:

None yet.

When were the last committers or PMC members elected?

None yet.

Signed-off-by:

[X](tamaya) John D. Ament
[X](tamaya) Mark Struberg
[](tamaya) Gerhard Petracek
[](tamaya) David Blevins

Shepherd/Mentor notes:

John D. Ament (johndament):

Podling is struggling a bit with some growing pains. Hope it will
straighten out soon.

Taverna

Taverna is a domain-independent suite of tools used to design and execute
data-driven workflows.

Taverna has been incubating since 2014-10-20.

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

1. IP clearance to be signed by Univ of Manchester
2. Move source code, wiki, issues to apache.org.
3. Re-awake community after migration, prepare a 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?

Quiet month - dominated by migration issues.

Involvement with students who are combining Taverna and Apache Storm.

Email traffic December 2014:

* dev@taverna: 50 messages
* users@taverna: 12 messages

How has the project developed since the last report?

The project is waiting for the Software Grant from Univ of Manchester
Solicitor and Contracts Manager as he needed to retrieve the existing CLAs
from archive. He has replied (2015-01-05): "They should arrive back from
storage this week and we can proceed with the [software grant] transfer"

Taverna's WSDL-support modified to work with Apache Woden.

About 45% of code base ready to be migrated (moved to slimmer git layout,
modified to org.apache.taverna groupId and build).
TODO: LICENSE and NOTICE plugins for Maven.

Apache Taverna website live at
<http://taverna.incubator.apache.org/>
-- not yet migrating content from
<http://www.taverna.org.uk/> until IP clearance sorted

Discussions around packaging / Debian

Date of last release:

None

When were the last committers or PMC members elected?

None since incubation

Signed-off-by:

[X](taverna) Andy Seaborne
[X](taverna) Chris Mattmann
[](taverna) Suresh Srinivas
[](taverna) Suresh Maru
[X](taverna) Marlon Pierce

Shepherd/Mentor notes:

Andy Seaborne (andy):

Given the podling has not yet got the software grant on ASF file, through no fault of the podling, I suggest keeping reporting Taverna monthly.

Justin Mclean (jmclean):

Off to a good start, with discussions about making builds and assembling LICENSE and NOTICE. Mentors (very) active.

Usergrid

Usergrid is Backend-as-a-Service (BaaS) composed of an integrated database (Cassandra), application layer and client tier with SDKs for developers.

Usergrid has been incubating since 2013-10-03.

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

1. Continue to build community
2. Perfect release process
3. Exclude unnecessary dependencies from release

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 added a couple of new committers: Alex Muramoto, Ryan Bridges and John Ament. We are working on our second release from the Incubator, Usergrid 1.0.1 and are creating release candidates now for voting.

How has the project developed since the last report?

Much work has been done in the "two-dot-o" version of Usergrid with a new persistence and query engine. Bugs have been fixed in the master branch and it is now being prepared for a Usergrid 1.0.1 release.

Date of last release:

Back in September: 2014-09-10

When were the last committers or PMC members elected?

Just last week - 2014-12-19

Signed-off-by:

[X](usergrid) Dave Johnson
[X](usergrid) Jake Farrell
[X](usergrid) Jim Jagielski

[X](usergrid) John D. Ament
[](usergrid) Lewis John Mcgibbney
[](usergrid) Luciano Resende

Shepherd/Mentor notes:

John D. Ament (johndament):

The group is still getting accustomed to the Apache Way. Special thanks to those in the incubator who have helped with licensing issues in the current release attempt.

Zeppelin

Zeppelin is a collaborative data analytics and visualization tool for distributed, general-purpose data processing systems such as Apache Spark, Apache Flink, etc.

Zeppelin has been incubating since 2014-12-23.

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

1. Complete the SGA
2. Migrate to ASF infrastructure
3. Complete PODLINGNAMESEARCH-64

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?

Zeppelin has just entered the ASF Incubator and the onboarding is progressing nicely. All initial committers have been issued ASF accounts and the basic INFRA has been setup.

How has the project developed since the last report?

This is the first report.

Date of last release:

N/A

When were the last committers or PMC members elected?

N/A

Signed-off-by:

[x](zeppelin) Konstantin Boudnik
[X](zeppelin) Hyunsik Choi
[X](zeppelin) Ted Dunning
[X](zeppelin) Henry Saputra
[X](zeppelin) Roman Shaposhnik