# October2014

/!\ *FINAL* (!)

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

---

## Incubator PMC report for October 2014

### Timeline

| | |
|---|---|
| Wed October 01 | Podling reports due by end of day |
| Sun October 05 | Shepherd reviews due by end of day |
| Sun October 05 | Summary due by end of day |
| Tue October 07 | Mentor signoff due by end of day |
| Wed October 08 | Report submitted to Board |
| Wed October 15 | Board meeting |

### Shepherd Assignments

| | |
|---|---|
| Alan D. Cabrera | Parquet |
| Andrei Savu | Ripple |
| Dave Fisher | ODF Toolkit |
| John Ament | Aurora |
| Justin Mclean | DataFu |
| Konstantin Boudnik | Samza |
| Matthew Franklin | Argus |
| Matthew Franklin | Optiq |
| Raphael Bircher | MetaModel |
| Roman Shaposhnik | REEF |
| Roman Shaposhnik | Usergrid |
| Ross Gardler | DeviceMap |
| Suresh Marru | Streams |

### Report content

```
Incubator PMC report for October 2014

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 32 podlings under incubation this month. No projects are
currently voting on graduation. The incubator added Ignite as a new
incubating project. DeviceMap, Drill, Falcon, Flink, Johnzon, Usergrid
all had releases since the last report. DeviceMap is ready to graduate.

* Community

  New IPMC members:

    No additions
```

People who left the IPMC:

         (none)


* New Podlings

     Ignite

* Graduations

     The board has motions for the following:

         (none)

* Releases

     The following releases were made since the last Incubator report:

     Apache DeviceMap Data 1.0.1 incubating
     Apache Drill 0.5.0-incubating
     Apache Falcon 0.5-incubating
     Apache Flink 0.6.1-incubating
     Apache Johnzon 0.1-incubating
     Apache Usergrid 1.0


* IP Clearance

         FlatSpark for Apache Flex

* Legal / Trademarks

         None

* Infrastructure

         Nothing to report

* Miscellaneous

     Apache Optiq becomes Apache Calcite

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

* Still getting started at the Incubator

     REEF

* Not yet ready to graduate

     No release:

         Argus
         DataFu
         Parquet
         Streams

     Community growth:

         Aurora
         MetaModel
         Optiq
         Samza
         Usergrid

* Ready to graduate

         DeviceMap

* Did not report, expected next month

  ODF Toolkit
  Ripple


--------------------------------------------------------------------
                        Table of Contents

------------------------------------------------------------------


--------------------
Argus

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

Argus has been incubating since 2014-07-24.

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

  * Get approval for name of Argus (PODLINGNAMESEARCH-54)
      * Requested a name change to "Apache Ranger"
  * Make a release
  * Add new committers

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

How has the project developed since the last report?

  * Create a project website
  * Working towards first incubator release (version 0.4.0)
  * 54 jiras resolved this month

Date of last release:

  none yet

When were the last committers or PMC members elected?

  none yet

Signed-off-by:

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


--------------------
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. Second release, progress being tracked in ticket AURORA-711
  2. Establish a consistent development process and release rhythm.
  3. Expanding the community and continuing to add new committers.

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 addition:   David McLaughlin, 2014-08-20

  Issue backlog status since last report:

  * Created:   176
  * Resolved:  151

  Mailing list activity since last report:

  * @dev       340 messages
  * @issues    774 messages
  * @reviews  1402 messages

How has the project developed since the last report?

  - IRC weekly meetings, digests sent to dev@ list
  - Automated release workflows
  - Fixing issues brought up in our first release

Date of last release:

  - Apache Aurora 0.5.0 (incubating) - 2014-08-01


Signed-off-by:

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

--------------------
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. Building an ASF-based community.
  2. Release.
  3. Adding support for Hadoop 2.x

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?

   Three new users have contributed code since the last report.

How has the project developed since the last report?

   A couple more UDFs have been committed.  One bug fix was committed.  All JARs have been removed from the repo
(a blocker for source release).
   A build task has been added for creating a source release.  No open blockers for release left at this point.
   Several more UDFs have been contributed but are still under review.

Date of last release:

   No release yet.

When were the last committers or PMC members elected?

   2014-02-22

Signed-off-by:

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

Shepherd/Mentor notes:

(jmclean): Did not report on time. Low level mentor activity but no obvious issues other than missing release.
(Release mentioned in last report and DATAFU-53 blocking release has been resolved).


--------------------
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.

The mentors feel DeviceMap is ready to graduate, the new PMC's roster has been discussed and
the graduation resolution is being prepared.

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 change. The DeviceMap community will probably remain small, graduating with 5 initial PMC members.

How has the project developed since the last report?

The project has started its incremental device data release with its recent 1.0.1 release.
The project also is planning to release its new 1.1 clients (Java and .NET). Long term plans
have been made to overhaul the device data in the 2.0 release.

Date of last release:

   2014-09-29

When were the last committers or PMC members elected?

May 17th, 2013

Signed-off-by:
```

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


   --------------------
   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. More contributors and committers to increase diversity of the community
      2. More releases under ASF
      3. Connect with other projects from Apache's ecosystem such as Apache Gora, Apache Flink

   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?

      A few new people have started to work actively on the mailing list
      and contributing patches and new features.

      One of the new contributors, Alberto Rodriguez, had added new module to
      connect to ElasticSearch.

      Updated ASF incubator page to reflect current state.

      Updated MetaModel website to contain link on how to contribute.

   How has the project developed since the last report?

      Apache MetaModel 4.2.0-incubating was released on July 28, 2014.

      We are near of proposing to have new committer.

      Mailing list activity since last report:

         @dev via http://mail-archives.apache.org/mod_mbox/incubator-metamodel-dev/
            September 137 messages
            August 28 messages
            July 109 messages

            51 members

         @commits via http://mail-archives.apache.org/mod_mbox/incubator-metamodel-commits
            September 26 messages
            August 2014Thread 3 messages
            July 2014Thread 47 messages

   Date of last release:

      July 28, 2014

   When were the last committers or PMC members elected?

      June 30, 2014


   Signed-off-by:

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

```
--------------------
Optiq (currently renaming to Calcite)

Optiq/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.

Optiq/Calcite has been incubating since 2014-05-19.

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

  1. Complete renaming to Calcite.

  2. Re-organize code into org.apache.calcite namespace

  3. Create a product worthy of a "1.0" label.

  4. Improve web site

  5. Build an ASF 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 held our first meetup, a hackathon hosted by MapR. It was well-attended,
  and we attracted a number of new contributors at various skill levels,
  who produced a number of contributions that have since been committed.

  Julian Hyde spoke to over 200 attendees at the Seattle Scalability meetup
  on how Apache Hive uses Apache Optiq to do its cost-based optimization.

  In September, there were 72 messages on dev@, 38 commits, 36 issues created,
  30 issues resolved.

How has the project developed since the last report?

  We completed https://issues.apache.org/jira/browse/PODLINGNAMESEARCH-51
  and concluded that Optiq was not a viable name.
  We agreed (via PPMC vote and consensus on dev) that Calcite was suitable new
  name, and logged https://issues.apache.org/jira/browse/PODLINGNAMESEARCH-56,
  which has been approved by legal. Now we are part way through renaming;
  see https://issues.apache.org/jira/browse/INFRA-8413

  We expect to have a release candidate for apache-calcite-0.9.1-incubating
  to go to a PPMC vote this week.

Date of last release:

  2014-08-25

When were the last committers or PMC members elected?

  2014-07-15

Signed-off-by:

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

--------------------
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
  - 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:         N/A
  * Contributor addition:  N/A

Issue backlog status since last report:

  * Created:   27
  * Resolved:  19

Mailing list activity since last report:

  * dev      144 messages

How has the project developed since the last report?

  - Attempted parquet-format release twice, next RC in early October.
  - Assembled tasks to complete for a parquet-mr release
  - New push-down filter API and task-side block metadata reading

Date of last release:

  - No releases as of yet.

Signed-off-by:

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


--------------------
REEF

REEF (Retainable Evaluator Execution Framework) is a scale-out computing
fabric that eases the development of Big Data applications on top of resource
managers such as Apache YARN and Mesos.

REEF has been incubating since 2014-08-12.

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

  1. Get the code transferred from a Microsoft-owned GitHub account to Apache infra.
  2. Line up the Microsoft internal process and approvals for (2)
  3. Update the existing project web site and figure out how to transfer it to the ASF infrastructure.

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?

  * All committers have setup @apache.org aliases and JIRA accounts.
  * Replaced the Google Groups based mailing list with an ASF hosted list for all dev communications
```

How has the project developed since the last report?

   * Issue tracking is beginning to move from GitHub to JIRA.
   * The name `Apache REEF` has been approved (PODLINGNAMESEARCH-55)
   * Microsoft-Internal approval for the code transfer and donation are there


Date of last release:

   None yet.

When were the last committers or PMC members elected?

   None Yet

Signed-off-by:

   [X](reef) Chris Douglas
   [ ](reef) Chris Mattmann
   [ ](reef) Ross Gardler
   [ ](reef) Owen O'Malley

--------------------
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:

   1. Finish outstanding 0.8 bugs, so we can do our second incubator release.
   2.
   3.

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?

Very well. Numerous presentations at conferences. Several instances of meetups/presentations by people outside
the committer community. Mailing list traffic has been slow but steady.

How has the project developed since the last report?

Very well. We've closed off all of our major issues for our 0.8 release, save one. Once the last issue is
finished (in progress now), we're going to do a release, and then begin the incubation graduation process.

Date of last release:

   2014-06-27

When were the last committers or PMC members elected?

   2014-08-13

Signed-off-by:

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

--------------------
Streams

Apache Streams is a lightweight server for ActivityStreams.

Streams has been incubating since 2012-11-20.

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

  1. Community Growth
  2. PPMC picks up the duties required of it, rather than rely on the mentors or let them languish
  3. Generate Documentation

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

N/A

How has the community developed since the last report?

The community engagement has been strong for this period. The list serve has had multiple discussions around
project standards and generating documentation.  Also the number of community members comment on pull requests
has increased.

How has the project developed since the last report?

20 pull requests have been open in Sept.
15 pull requests have been merged in Sept.
55 comments have been made on those pull requests
6 community members have opened pull requests in Sept.

Date of last release:

  2013-01-09

When were the last committers or PMC members elected?

2014-01-14 (Carol Hansen)

Signed-off-by:

  [x](streams) Matt Franklin
  [x](streams) Ate Douma
  [ ](streams) Craig McClanahan


--------------------
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. Make another incubating release (most likely 1.1)
  2. Continue to grow the Usergrid community
  3. Document and streamline the Usergrid release process

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?
The community continues to develop in a positive fashion.
Usergrid was represented (indirectly) at this years CassandraSummit in San Francisco.
Various mentions of UG at that event, amongst the other community events which the UG
PPMC put together are building the user base.
The #usergrid IRC channel continues to see people come and go with people being directed
to the mailing lists whenever issues cannot be addressed in a more direct/instant manner.

user@ Traffic
Sep 2014 --> 7
Aug 2014 --> 13
Jul 2014 --> 9

```
dev@
Oct 2014 --> 5
Sep 2014 --> 127
Aug 2014 --> 288
Jul 2014 --> 144

commits@
Oct 2014 --> 101
Sep 2014 --> 443
Aug 2014 --> 764
Jul 2014 --> 136
```

How has the project developed since the last report?
The project has come on leaps and bounds post release of Usergrid 1.0 (incubating).
The development (PPMC) team worked very well together, and with the IPMC to ensure that
the 1.0 release went smoothly. The release procedure was slightly longer than we would have
liked however lessons have been learned and positives have been taken from the process as a whole.
Procedural nuances have been ironed out and we look forward to a 1.1 release in the near future.
In addition we have been documenting material on the Usergrid Incubation status page.
http://incubator.apache.org/projects/usergrid.html
Final note, KUDOS to Rod and Dave from the Usergrid PPMC who acted as release managers.

Date of last release:

  2014-09-10

When were the last committers or PMC members elected?

2014-07-17 New committer: George Reyes

Signed-off-by:

  [ ](usergrid) Dave Johnson
  [X](usergrid) Jake Farrell
  [ ](usergrid) Jim Jagielski
  [X](usergrid) Lewis John Mcgibbney
  [ ](usergrid) Luciano Resende