

# June2014

/!\ \*FINAL\* (!)

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

---

## Incubator PMC report for June 2014

### Timeline

Wed June 04	Podling reports due by end of day
Sun June 08	Shepherd reviews due by end of day
Sun June 08	Summary due by end of day
Tue June 10	Mentor signoff due by end of day
Wed June 11	Report submitted to Board
Wed June 18	Board meeting

### Shepherd Assignments

Alan D. Cabrera	Brooklyn
Andrei Savu	Slider
Andrei Savu	Streams
Dave Fisher	S4
John Ament	Storm
Justin Mclean	Kalumet
Matthew Franklin	Falcon
Konstantin Boudnik	Flink
Matthew Franklin	MRQL
Raphael Bircher	Wave
Konstantin Boudnik	Drill
Ross Gardler	DeviceMa p
Suresh Marru	Sentry
Suresh Marru	log4cxx2
Roman Shaposhnik	NPanday

### Report content

Incubator PMC report for June 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 currently 34 podlings under incubation. Two new podlings joined the Apache Incubator in May, two podlings graduated.

\* Community

New IPMC members:

John D. Ament

People who left the IPMC:

(none)

\* New Podlings

Optiq  
Parquet

\* Graduations

Phoenix  
Stratos

The board has motions for the following:

(none)

\* Releases

The following releases were made since the last Incubator report:

metamodel-4.1.0-RC1-incubating  
odftoolkit-0.6.1-incubating  
sentry-1.3.0-incubating  
slider-0.30-incubating  
spark-0.8.1-incubating

ODF Toolkit's release was approved using the alternative voting process worked out last November/December. We believe that's the first official release in the history of the Foundation that passed with binding votes by non-PMC members.

\* IP Clearance

Cray donated a set of libraries called Hadoop RDF Tools, and a VOTE to accept the tools was called by Rob Vesse. The vote passed by lazy consensus after 72 hours. No objections or points of discussion were raised during the VOTE.

\* Legal / Trademarks

New name for Stratosphere, now Flink.

\* Infrastructure

Apache Podlings affected by ASF mail outage from May 7th through May 10th, and extended recovery period to deliver backlog.

\* Miscellaneous

S4 failed to report this month, and has been generally inactive. It is likely that S4 will be retired, though the PPMC has not voted yet to do so.

NPanday failed to report in a previous cycle, but was accidentally dropped from the reporting schedule for this month. NPanday developers and mentors were contacted. In 72 hours nobody was available to report on the project. Given consistent troubles with getting project to report and project mentors MiA, IPMC will look into a resolution.

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

\* Still getting started at the Incubator

Brooklyn  
Parquet

\* Not yet ready to graduate

No release:

DeviceMap  
Flink  
log4cxx2  
Wave

Community growth:

Falcon  
MRQL  
Sentry  
Slider  
Streams

\* Ready to graduate

Drill  
Storm

\* Did not report, expected next month

Kalumet  
S4  
NPanday

---

#### Table of Contents

Brooklyn  
DeviceMap  
Drill  
Falcon  
Flink (formerly Stratosphere)  
log4cxx2  
Kalumet  
MRQL  
Parquet  
Sentry  
Slider  
Storm  
Streams  
Wave

---

Brooklyn

Brooklyn is a framework for modeling, 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. Migrating fully to Apache infrastructure (mainly concerning open pull requests and issue tracker)
2. Forming a diverse community and PPMC
3. Learning and following "the Apache way"

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?

This is our first report; our community remains very similar to the point where we joined the Incubator.

How has the project developed since the last report?

This is our first report; our project remains very similar to the point where we joined the Incubator.

Date of last release:

No releases yet under the Incubator.

When were the last committers or PMC members elected?

No new committers since we joined the Incubator.

-----  
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.

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

1. Grow community.
2. Generate a release.
3. Explore options to allow contributions through a Web based interface/API to add new device specs.
4. Improved support for additional languages/platform (.NET, PHP,...)

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?

Tangentially, there have been presentations at conferences like JavaLand raised awareness. A Mobile Development MeetUp group in Germany is looking into the code, and where they might contribute.

How has the project developed since the last report?

Syncing device data with OpenDDR in regular intervals when relevant changes occur. Update of the Java DDR Simple API and other libraries.

Date of last release:

N/A

When were the last committers or PMC members elected?

N/A

-----  
Drill

Description:

Apache Drill is a distributed system for interactive analysis of large-scale datasets that is based on Google's Dremel. Its goal is to efficiently process nested data, scale to 10,000 servers or more and to be able to process petabytes of data and trillions of records in seconds.

Drill has been incubating since 2012-08-11.

Three Issues to Address in Move to Graduation:

1. Complete the 1.0 feature set (team targets next release and graduation in the month of July)

2. Continue to attract new developers/contributors with a variety of skills and viewpoints
3. Continue the outreach activities to build the early user community for the technology

Issues to Call to Attention of PMC or ASF Board:

None

How community has developed since last report:

Community awareness and outreach were strengthened in multiple forums as below

- \* First Apache Drill Hackathon was organized on 4/24. Over 40 participants including members from Visa, LinkedIn, Cisco, Hortonworks worked to harden/enhance Drill project. Several new features have been added to Drill product Array reference functions, enhanced Optiq support, Kafka storage plugin, robust testing framework etc
- \* Hive big data think tank meet up on 5/14- Talk by MC Srivas, with ~200 member participation
- \* Open Source Cloud meet up on 4/23 - Talk by Keys Botzum
- \* Apache Conference session on 4/8 - Talk by Neeraja Rentachintala, with ~100 members participation

Apache Drill is also showcased at the Hadoop Summit 6/3-6/5

Mailing list discussions:

Activity summary for the user mailing list:

[http://mail-archives.apache.org/mod\\_mbox/incubator-drill-user/](http://mail-archives.apache.org/mod_mbox/incubator-drill-user/)

- \* June to date 6/10: 28
- \* May 2014, 82
- \* March 2013, 15

Activity summary for the dev mailing list:

[http://mail-archives.apache.org/mod\\_mbox/incubator-drill-dev/](http://mail-archives.apache.org/mod_mbox/incubator-drill-dev/)

- \* June to date 6/10: 87 (jira focussed discussions were removed from this thread recently)
- \* May 2014, 1183(jira, focused discussions)
- \* April 2014, 772 (jira; focused discussions)

For details of code commits, see

<https://github.com/apache/incubator-drill/graphs/commit-activity> (about 300 commits in the past 3 months)

26 contributors have participated in GitHub code activity; there have been 142 forks.

Community Interactions

Weekly Drill hangout continues, conducted remotely through Google hangouts Tuesday mornings 10am Pacific Time to keep core developers in contact in realtime despite geographical separation.

Community stays in touch through @ApacheDrill Twitter ID, and by postings on various blogs including Apache Drill User <http://drill-user.org/> which has had several updates and through international presentations at conferences.

Articles

Examples of articles or reports on Apache Drill since last report include:

- \* Drill Hackathon summary blog post by Jacques Nadeau
- \* Drill milestone roadmap blog post by Neeraja Rentachintala

\* Drill code samples by Nitin Bandugula

Social Networking

@ApacheDrill Twitter entity is active and has grown substantially by 19%, to 887 followers.

How project has developed since last report

Significant progress is being made on the performance and distributed optimization

C++ client API and ODBC driver leveraging the C++ API was built for Drill by a group led by George Chow in Vancouver. The initial drops for the driver are available

New functionality has been added to the product namely distributed optimization, join order optimization, Table/view creation, repeated map support, HBase support, expanded SQL support, Text readers, new data types and functions, Session options for query tuning and lot more

Nearly ~500 bugs files and ~400 bugs resolved

Significant progress on running ANSI standard queries such as TPC-H

Significant code drops have been checked in from a number of contributors and committers

New docs have been published on Drill wiki (Apache Drill in 10 mins, Working with various data sources and Installing and Running Apache Drill on a cluster)

Work toward a Beta milestone is progressing substantially.

Signed-off-by:

[x](drill) Ted Dunning  
[x](drill) Grant Ingersoll  
[ ](drill) Isabel Drost-Fromm  
[x](drill) Sebastian Schelter

Shepherd/Mentor notes:

Konstantin Boudnik (cos):

Project's dev@ list is very active both with the JIRA traffic and otherwise. June report to the board hasn't been sent on time.

-----  
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. Continue to build a broader community
2. Add more committers 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?

\* More users & contributors have joined the falcon project and the

community continues to grow rapidly

- \* We have had 2 Birds of Feather for Falcon at Hadoop Summit, Europe and USA. There were quite a few registered users

How has the project developed since the last report?

- \* 0.5-incubating release candidate is available for vote
- \* Project now has diverse user & contributor base
- \* It appears to be time for the community to start thinking about graduation

Date of last release:

- \* 2014-02-03 (0.4-incubating)

When were the last committers or PMC members elected?

- \* Two new committers have joined the project on 2014-03-07

Signed-off-by:

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

-----  
Flink

Flink is an open source system for expressive, declarative, fast, and efficient data analysis. Flink combines the scalability and programming flexibility of distributed MapReduce-like platforms with the efficiency, out-of-core execution, and query optimization capabilities found in parallel databases.

Flink has been incubating since 2014-04-14.

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

1. Infrastructure setup (git, JIRA, mailing list renaming)
2. Release
3. 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?

The community still "lives" outside the ASF since the infrastructure is not ready.

Two new contributors opened pull requests since the last report.

How has the project developed since the last report?

The community has voted for a new name, so the project is moving away from its old name "Stratosphere" to the new one: "Flink".

We are still waiting for Infra to rename the mailing list, to create a JIRA (blocked by the mailing list) and create a Git repository (blocked by the SGA (which we already provided)).

We released a new major release (0.5) outside the ASF.

Date of last release:

2014-05-31 (no incubator release yet)

When were the last committers or PMC members elected?

Only initial committers.

Signed-off-by:

[ ](flink) Sean Owen  
[X](flink) Ted Dunning  
[ ](flink) Owen O'Malley  
[X](flink) Henry Saputra  
[ ](flink) Ashutosh Chauhan  
[X](flink) Alan Gates

Shepherd/Mentor notes:

Konstantin Boudnik (cos):

Project dev@ list has healthy traffic. Some of the mentors are active on the list and helping community. No visible issues.

-----  
log4cxx2

Logging for C++

log4cxx2 has been incubating since 2013-12-09.

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

1. We probably need to decide what's an acceptable platform/compiler list, sort out any remaining hiccups, and then move towards a release.
2. We need to release 0.11.0 as maintenance release done via ASF
3. We need to make a broader use of the ASF infrastructure (notably the CI, as well as improved management of the Jira site - notably with triaging), and establish a roadmap for the next releases.

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?

No changes in team.

How has the project developed since the last report?

We are actively supporting the project. Mails suggest that there's work going on in the background by some committers especially regarding the build process, but there hasn't been any notable commit in the last weeks. The current trunk now fixes most of the important outstanding bugs.

Date of last release:

2008-04-03 was the official, pre-incubation 0.10.0

Many post-0.10.0 commits exist on trunk which we intend to get out as 0.11.x.

When were the last committers or PMC members elected?

N/A

Signed-off-by:

[X](log4cxx2) Christian Grobmeier  
[ ](log4cxx2) Scott Deboy

-----  
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.

No report submitted this month.

Shepherd/Mentor notes:

Justin Mclean (jmclean):

Podling may be in a little trouble. Failed to report this month, mailing list traffic and commits at a very low level. A mentor is still active but looks like the last release candidate (October 2013) wasn't voted on by the IPMC and a new release candidate hasn't been created. Last report talked about graduation so perhaps they just need some help in getting past the final hurdle?

-----  
MRQL

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

MRQL has been incubating since 2013-03-13.

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

1. increase adoption, expand user community, and increase user list activity
2. have at least one more incubator 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?

A new PMC member was elected, Moon Soo Lee. The company DataSayer and the project Zeppelin (zeppelin-project.org) are now using MRQL.

How has the project developed since the last report?

We have switched to Junit for query testing, instead of evaluating queries from files using plain Java code. New tests were introduced and some bugs were corrected based on these tests.

Date of last release:

2013-10-31

When were the last committers or PMC members elected?

2014-04-17

Signed-off-by:

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

-----  
Parquet

Parquet is a columnar storage format for Hadoop.

Parquet has been incubating since 2014-05-20 .

Three most important issues

- Finish bootstrapping project, IP clearance, initial website
- Expanding the community and adding new committers
- 1st 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?

- All initial committers have submitted ICLAs and the accounts have been created. The mailing lists have been setup and we are starting to use them for communication.

How has the project developed since the last report?

- We have setup the incubator status page and are waiting on the final SGA to be sent in to start the code import (INFRA-7782)

Date of last release

- No releases as of yet. Working through initial IP clearance.

When were the last committers or PMC members elected?

- N/A, still bootstrapping the project.

Signed-off-by:

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

-----  
S4

S4 (Simple Scalable Streaming System) is a general-purpose, distributed, scalable, partially fault-tolerant, pluggable platform that allows programmers to easily develop applications for processing continuous, unbounded streams of data.

S4 has been incubating since 2011-09-26.

No report submitted this month. There have been discussions about retiring the podling from the Apache Incubator.

-----  
Sentry

Sentry is a highly modular system for providing fine grained role based authorization to both data and metadata stored on an Apache Hadoop cluster.

Sentry has been incubating since 2013-08-08.

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

1. Grow the Sentry community
2. Add new committers that diversify the project
3. Continue to release at regular intervals

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 is growing. A number of new members have started actively contributing to the project.

How has the project developed since the last report?

The project did its 2nd release from incubation. Thanks to lot of hard work from the release manage Karthik Ramchandran and other community members, version 1.3 was released earlier this month. The db backed policy store feature branch is merged into master. The community is working towards improving that functionality and stabilization.

Date of last release:

2014-05-15

When were the last committers or PMC members elected?

Vamsee Yarlagadda has been added as a committer as announced on Feb 28, 2014. No new PPMC members have been added since the project has entered the incubator.

Signed-off-by:

[X](sentry) Arvind Prabhakar  
[X](sentry) Joe Brockmeier  
[X](sentry) David Nalley  
[ ](sentry) Olivier Lamy  
[X](sentry) Patrick Hunt  
[X](sentry) Thomas White

-----  
Slider

Slider is a collection of tools and technologies to package, deploy, and manage long running applications on Apache Hadoop YARN clusters.

Slider has been incubating since 2014-04-29.

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

1. Podling name search
2. Growing user community
3. Growing and increasing diversity of developer 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're only getting started, and do not have any diversity yet. Our first goal will be to get users, bug reporters and developers. The dev list is up and running -and we do not have any separate user list, so we hope to pull users into coding.

The initial release will help to gain awareness, and should bring in users.

How has the project developed since the last report?

Our main activity has been setting up the incubator-hosted project, JIRA, site, mailing list. etc. The Infra team have been very helpful here.

- Making and publishing our site has educated us on site publishing
- The initial release has taught us of the release process for incubating projects, and where we need to improve it.

We're still trying to get Jenkins building, but that's a matter of handling protobuf-versions on the build machines, rather than infra

involvement.

We have made our first ASF-hosted release on June 2. This was a source only release; some of the reviewers have made suggestions which will need to be incorporated into the next. We plan to do another release in a few weeks time, and so evolve the project rapidly.

Date of last release:

2014-06-02

When were the last committers or PMC members elected?

We have only just set up and so the committer and PMC membership is as covered in the incubator proposal.

Signed-off-by:

[ ](slider) Arun C Murthy  
[ ](slider) Devaraj Das  
[ ](slider) Jean-Baptiste Onofré  
[X](slider) Mahadev Konar

-----  
Storm

Storm is a distributed, fault-tolerant, and high-performance realtime computation system that provides strong guarantees on the processing of data.

Storm has been incubating since 2013-09-18.

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

1. Community growth
2. Next release
3. Documentation improvements

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 have added on additional committer and community activity in the mailing lists continues to grow.

How has the project developed since the last report?

We have held a logo contest and completed the first of two rounds of voting. Final logo selection will take place shortly. We have also established a feature branch for incorporating all the security work done by the engineering team at Yahoo!. Having the option for secure Storm deployments will play a big role in increasing enterprise adoption.

We plan on initiating our next release next week in addition to completing the logo contest. Once these two tasks are complete we would like to consider beginning the graduation process.

Date of last release:

2014-02-22

When were the last committers or PMC members elected?

2014-03-26 (Michael G. Noll)

Signed-off-by:

[x](storm) Ted Dunning

[X](storm) Arvind Prabhakar  
[ ](storm) Devaraj Das  
[ ](storm) Matt Franklin  
[ ](storm) Benjamin Hindman

Shepherd/Mentor notes:

John Ament (johndament):

Storm is showing a lot of progress and growth. They have a lot of activity on JIRA and mailing lists, and are generally in good shape.

-----  
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. Growing the community
2. Consolidate codebase and produce a release
3. Continue to implement the Apache way and community decision making

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

None

How has the community developed since the last report?

There are contributions from new individuals and an increase in new faces on the dev list.

How has the project developed since the last report?

There has been a significant increase in development and communication on the dev list; but, the community still needs to grow in size and adoption of the Apache Way.

Date of last release:

2013-01-09

When were the last committers or PMC members elected?

2014-01-14 (Carol Hansen)

Signed-off-by:

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

-----  
Wave

A wave is a hosted, live, concurrent data structure for rich communication. It can be used like email, chat, or a document.

Wave has been incubating since 2010-12-04.

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

1. Make a release
2. Expand the community
3. Decide on a future path (whether we are continuing?)

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

None.

How has the community developed since the last report?

There has been no major development. A new user reported a regression on master compared to the last release candidate, but overall the project has been very quiet this last quarter.

How has the project developed since the last report?

Some maintenance work has been committed from new developers, this replaced Socket.IO with Atmosphere. A few other minor fixes for different aspects (gadgets, build system) have been committed.

Date of last release:

N/A

When were the last committers or PMC members elected?

30th June 2013.

Signed-off-by:

☒ (wave) Christian Grobmeier  
☐ (wave) Upayavira