October2011

October 2011 Board reports (see ReportingSchedule).

These reports were due here by **Wednesday, 12 October 2011** so that the Incubator PMC could relay them to the board.

This report is closed.

Your project might need to report even if it is not listed below, please check your own reporting schedule or exceptions.

Please remember to include:

- The "incubating since" info.
- The project's top 2 or 3 things to resolve prior to graduation.
- A short description of what your project's software does.
- The **Signed off by mentor:** is for Mentor(s) to show that the Report has been reviewed.

Remember that this is a "text" report, so no markup please.

```
--------------------
# Accumulo

Accumulo is a sorted, distributed key/value store based on BigTable's design.  Accumulo entered incubation in
September 2011.

In the move towards graduation, we must address:
1. Learning Apache procedures
2. Creating releases
3. Building a community

Any issues that the Incubator PMC or ASF Board might wish/need to be aware of:
Discussion is ongoing as to whether agreements other than ICLAs are needed or desirable for people employed by
the US government to make contributions to Apache (see LEGAL-100).

Developments since entering incubation:
* mailing lists created
* JIRA created
* SVN directory and git mirror created
* accounts for initial committers created
* ICLAs and Software Grant filed
* CMS-ready site begun
* initial code uploaded
* Jenkins build created
* ReviewBoard group created

Signed off by mentor:
Alan Cabrera

--------------------
# ACE

Apache ACE is a software distribution framework that allows you to centrally manage and distribute software
components, configuration data and other artifacts to target systems. ACE started incubation on April 24th 2009.

There are currently no issues requiring board or Incubator PMC attention.

Community:
 * We've got a lot of great feedback and patches from the community.
 * There have been talks with the jclouds (we use them) as well as the Amdatu (they use us) open source
projects.

Software:
 * We now have a REST client API.
 * The management agent has been extended.
 * Karaf features were added.
 * We have a server side resolver based on Apache Felix.

Licensing and other issues:
 * None at the moment.

Things to resolve prior to graduation:
```

* We hope this is our very last board report as we think we're ready for graduation now!

Signed off by mentor: cziegeler

Ambari

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

 * Incubating since 30 August 2011.
 * Mailing lists created and mentors subscribed.
 * Confluence created.
 * Initial code committed.
 * Site created.
 * Code grant received.
 * Development proceeding actively.
 * RAT added to pom and report is clean.

Signed off by mentor: omalley

Any23

Anything To Triples (shortly Any23) defined as a Java library, a Web service and a set of command line tools to extract and validate structured data in RDF format from a variety of Web documents and markup formats. Any23 is what it is informally named an RDF Distiller.

A list of the three most important issues to address in the move towards graduation

    1. Port Any23 code to ASF infrastructure and update license headers
    2. Develop a strong community with organizational diversity and with strong connections to other relevant ASF communities.
    3. At least one Any23 incubating release

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

No, not at this time.

How has the community developed since the last report?

Everyone has their ASF account set up, except for Giovanni Tumarello (ICLA filed and Chris working on getting account set up). We've already moved our discussions onto the public mailing lists now that they are set up.

How has the project developed since the last report?

Any23 was voted into the Incubator by the IPMC on October 1, 2011.

Software grants from DERI and from FBK have been filed and accepted thanks to Michele Mostarda. There has already
been discussion of how to move over JIRA issues and Wiki issues from the Google Code site and the full issues transition
was completed on 12/10/2011. Code imports will likely begin soon, and a pointer on the existing Google Code site will point to the ASF Incubator
as the new home. Simone has created us a podling status page. Chris filed INFRA-3978 to track JIRA, Wiki and mailing list creation for podling set up. It's all taken care of (thanks Gav).

Signed off by mentor: mattmann

Celix

Celix is an implementation of the OSGi Specification in C.

Celix entered incubation on November 2, 2010.

In the code not much has changed during the past few months, mostly due to vacations. Some (reported) bugs have been fixed, and some interesting examples have been added. These example where donated by a user, and follow the OSGi in Action book. This makes them a great addition to Celix.
Furthermore, some work has been done to make testing from CMake easier, and we are also looking into a method for mocking services to be able to test only one service without all its dependencies.

On the community site, Luminis and Thales are organizing an event (in Dutch) to introduce embedded/distributed companies to OSGi. During this event OSGi will be explained, Celix and PojoSR will be introduced. Also a small introduction in Apache and the "Apache Way" will be given, with the goal to invite people to use/join Celix. Also, in november a talk and demo will be given at the ApacheCon.

Most important issues are:

    Improve robustness (APR, error handling etc)
    Generate awareness and grow a community!

Signed off by mentor: marrs

Chukwa

Chukwa is Hadoop monitoring system

Chukwa is an open source data collection system for monitoring large distributed systems.

A list of the three most important issues to address in the move towards graduation

    1. Simplify the build process and provide sample to use cases.
    2. At least one Apache Incubator release.
    3. Develop a strong community with organizational diversity and well aligned with existing ASF projects as outlined here: http://wiki.apache.org/incubator/ChukwaProposal#Alignment

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

No, not at this time.

How has the community developed since the last report?

There are some activity on the user mailing list, and some new users participated in adopting Chukwa for processing and visualizing big data analytics.

How has the project developed since the last report?

There are several bug fixes, and migration to Hadoop Metrics 2 for Hadoop 0.20.203+.

Signed off by mentor: cdouglas

Deft

Deft is a non-blocking, asynchronous, event driven high performance web framework running on the JVM. The project entered incubation on 2011-07-08.

The most important issues to address in moving to graduation:

    1. Prepare an initial incubation release.
    2. Attract more committers, and other community members.
    3. Secure the previously used URL, "deftserver.org".

Any issues that the Incubator PMC (IPMC) or ASF Board wish/need to be aware of?
- Not at this time.

How has the community developed since the last report?
- There have been no significant developments.

How has the project developed since the last report?
- Outstanding tasks are still being worked on, and the branch has been moved to trunk.
- There is a vote pending on project name change, a task previously identified as needing to be addressed before graduation.

Signed off by mentor: ngn, Mohammad Nour (mnour)

Easyant

Easyant is a build system based on Apache Ant and Apache Ivy. Incubating since 2011-01-31.

Towards graduation, we need to:
* Creating a release
* Building a community

Since the last report:
* an initial committer of Easyant didn't provided its iCLA. Hence an iterative cleanup of the code has been done together with developers and mentors. This cleanup is finished, the IP clearance is considered done, acknowledge by mentors.
* no particular progress on the diversity of the community
* discussion has started to how setup the build which will fit the ASF release guidelines

Signed off by mentor: bodewig antoine

Empire-db

Apache Empire-db is a relational database abstraction layer that allows developers to take a more SQL-centric approach in application development than traditional ORM frameworks. Its focus is to allow highly efficient database operations in combination with a maximum of compile-time-safety and DBMS independence. Empire-db has entered Incubation in August 2008.

Activity since last report:
* a lot of work on the code has been done in order to address architectural changes that may significantly improve the ease-of-use and robustness of the code. These changes have been the biggest core code changes since incubation.
* Testing of the new features and preparation of a new release are in progress

Community
* many of our committers have worked and contributed to this upcoming release especially all new committers that have joined the project since incubation. This gives us confidence that community improvements are slow but steady and that once people have learned about the benefits of Empire-db they will stick to it.
* requests of new users on mailing lists have been answered and problems have been solved.

Top priorities prior to graduation:
* Finish and publish the release
* Go out and advertise the benefits of Empire-db in general and the advantages of the new features in particular

Signed off by mentor:

bimargulies@apache.org

Etch

(project add text here)

Signed off by mentor:

Giraph

Giraph is a large-scale, fault-tolerant, Bulk Synchronous Parallel
(BSP)-based graph processing framework that runs on Hadoop. Giraph
entered the incubator in August 2011.

Project developments:

* JIRA activity steadily increased (57 issues reported so far)
* added Jake Mannix and Dmitriy Ryaboy as committers/PPMC members

Next steps:
* Making a release.
* Update on 1 initial committer who still hasn't filed an ICLA. He
will fill off the ICLA by next week.

Signed off by mentor: mattmann, omalley

Hama

```
(project add text here)

Signed off by mentor:

--------------------
# HCatalog

(project add text here)

Signed off by mentor:
--------------------
# Kafka

(introduced to Apache incubator on Jul 4, 2011)

Kafka provides an extremely high throughput distributed publish/subscribe messaging system. Additionally, it
supports relatively long term persistence of messages to support a wide variety of consumers, partitioning of
the message stream across servers and consumers, and functionality for loading data into Apache Hadoop for
offline, batch processing.

A list of the most important issues to address in the move towards graduation

    1. Successful podling release.
    2. Invite diverse new active committers

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

Not at this time.

How has the community developed since the last report?

We are getting very close to release the very first version of Kafka in Apache. A couple of RCs have been
proposed and we are still fixing various license header issues reported by rat. The proposed RC fixed more than
40 issues. The main feature added in the RC is end-to-end compression support.

How has the project developed since the last report?

Activities in the mailing list remain high. Several patches from non-committers have been submitted, reviewed
and committed.

Signed off by mentor:
Alan Cabrera


--------------------
# Kalumet

*Apache Kalumet*

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

Kalumet was accepted into Apache Incubator on September 20, 2011. See [Vote|http://markmail.org/message
/2aadus43wut7e3pi]

Status Page information: [http://incubator.apache.org/projects/kalumet.html].

First report:
* Creation of mailing list.
* Creation of Jira project.
* Creation of Confluence Space.
* Various committers account creation.
* Pending: svn repository (svndump has been provided)

Issues before graduation
* Create Kalumet web site.
* Make an incubating release.
* Grow the community size and diversity.

Signed off by mentor: olamy,jbonofre
```

Kato

(project add text here)

Signed off by mentor:

Lucy

Lucy is a loose port of the Apache Lucene search engine library, written in C
and targeted at dynamic language users.

Lucy was voted into the Incubator on July 22, 2010.

Progress since the last report:

 * Released versions 0.2.0 and 0.2.1.
 * Added new committer Brad Harder.
 * Technical developments:
   * Eliminated one of two remaining dependencies covered by LEGAL-86
     (JSON::XS). Only Parse::RecDescent remains.
   * Reestablished Windows portability.
   * Continued Clownfish C porting.

Top priorities prior to graduation:

 * Disperse knowledge of codebase among developers.

Issues for Incubator PMC or ASF Board:

 * None at this time.

Signed off by mentor: mattmann

ManifoldCF

(project add text here)

Signed off by mentor:

Mesos

* Implementation of a replicated log has been completed (based on Paxos). Next steps will be to use the log to
enable the Mesos masters to have an accurate snapshot of the currently running frameworks, currently connected
slaves, etc. This will enable the Mesos master to be able to launch the schedulers rather than them needing to
be launched independently. In addition, the log will be useful to give to schedulers that want high-
availability.

* Numerous bugs have been fixed by new contributors: Vinod Kone, Brian Wickman, John Sirois.

* Numerous bug fixes and features were added surrounding the interaction Mesos has with ZooKeeper, including in-
process ZooKeeper testing (contributed by John Sirois).

* Progress has been made on moving the build system to autotools. We'd like to accomplish this before trying to
do a release.

Signed off by mentor:

ODFToolkit

ODFToolkit

The ODF Toolkit is a set of Java modules that allow programmatic creation, scanning and manipulation of
OpenDocument Format (ISO/IEC 26300 == ODF) documents. Unlike other approaches which rely on runtime
manipulation of heavy-weight editors via an automation interface, the ODF Toolkit is lightweight and ideal for

server use.

* ODF Toolkit entered incubation on Aug 1st, 2011.

* Most important issues to address.
  1) Growing the community, increasing diversity of committers
  2) Completing IP review and updating headers
  3) Successful podling release.

* Any issues that the Incubator PMC or ASF Board might wish/need to be aware of
  None at this time.

* How has the community developed since the last report
  The mailing lists are ready now for 55 days. We have 36 subscribers. We're starting to receive patches from
new contributors.

* How has the project developed since the last report.

Legacy Bugzilla issue tracker has been migrated to the Apache JIRA instance.   We're now working on the details
of source and binary release directory layouts and build automation to support the same in preparation for our
first release.

We're exploring enabling Jenkins support for the build.

Signed off by mentor:
yegor

Olio

Olio is a web 2.0 toolkit to help developers evaluate the suitability,
functionality and performance of various web technologies by implementing
a reasonably complex application in several different technologies.
Olio entered incubation in October 2008

There is no development activity on this project. There was at least 1 new user in the last quarter who used
Olio.

Graduation From Incubation:

The user base is not showing growth. Since this is a test application/workload,
it is being used as such for short-term projects.
I think the decision was made to simply keep the project so users can continue to access it.

Oozie

* Oozie is a workflow management and scheduler primarily for Hadoop based jobs.
* Oozie entered the incubation on July 11, 2011.

* A list of the three most important issues to address in the move towards graduation:

   * Make the first Oozie release from Apache incubation.
   * Improve the documentations: user, development for quicker adoption
   * Establish the formal contribution process (such as CTR vs RTC)

* Any issues that the Incubator PMC or ASF Board might wish/need to be aware of:

   * No issues.

* How has the community developed since the last report:
   * Oozie user entirely moved to Apache Incubator using new user group email for communication.
   * Oozie-user email list is active and nearly 75 emails were exchanged in last month.
   * oozie-dev email list had nearly 150 emails in last month.

* How has the project developed since the last report.

   * Development is completely based on Oozie code from Apache SVN.
   * Development activity became stronger during the last month. Nearly 12 people created 24 JIRA in last month.
   * Extensively using review-board for patches. Development community has addressed nearly 20 code reviews.

Signed off by mentor: Alan Gates (gates)

Signed off by mentor:

OpenOffice.org

* OpenOffice.org entered incubation 2011-06-13.

OpenOffice.org is an open-source, office-document productivity suite providing six productivity applications based around the OpenDocument Format (ODF). OpenOffice.org is released on multiple platforms. Its localizations support 110 languages worldwide.

* Most Important To Address:

1) Migration of the remaining legacy OpenOffice.org website's content and services to Apache infrastructure, including wiki, forums, mailing lists, and cross-service registration using customized software not already supported by Apache projects and infrastructure. Successful negotiation of governance migration of user-supported services brought under incubation. Resolution of copyright, license and notice for content miugrated from legacy OpenOffice.org website.

2) Completion of the IP-review portions of the incubation checklist, which will require getting an amended SGA from Oracle to cover additional source files; scrubbing of incompatible notices from SGA-licensed code and resolving provenance of other existing materials being migrated. There are serveral modules that have non-compliant dependencies which will need to be resolved.

3) A Successful Podling Release

* Issues for IPMC or ASF Board Awareness

None

* Community Development Progress

As of 2011-10-12 there are 76 committers with 55 on the PPMC, compared to 72 and 55 at last report.

To serve our Japanese contributor base we've added an ooo-general-ja list.

Active discussions with an autonomous group that runs a support bulletin board for OpenOffice on the possibility of that community joining the AOOo project.

Discussion of criteria for bringing on new committers and PPMC members, and whether or not there should be a separate criteria for each role, as well as how these decisions should be communicated to the project.

Discussion of roles for PPMC members including TDF liaison, branding, press, and legal affairs.

* Project Development Progress

Dev focus has been on IP review and replacement of copyleft dependencies.

Detailed planning continues on public wiki: https://cwiki.apache.org/confluence/display/OOOUSERS/

Migration of several parts of the OpenOffice.org infrastructure are in various stages.

Signed off by mentor: rgardler

RAT

(project add text here)

Signed off by mentor:

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 entered incubation on September 26th 2011.

* Issues before graduation
 1. Migrate code from github to Apache SVN and make an initial release on Apache with the current code (version 0.4 = version 0.3 + checkpointing + bugfixes)
 2. All committers to get ICLA
 3. Migrate web site from s4.io to Apache infrastructure
 4. Complete the current work on the new version (0.5), which is a major refactoring in order to simplify concepts, API, and introduce new features such as dynamic deployment and reliable channels
 5. Grow the community

Community
 1. Mailing lists have been created. We are already using them.
 2. Jira initialized and started to track issues
 3. Karma granted for committers with ICLA
 4. Confluence Wiki space created
 5. IRC channel set-up

Project developments
 1. We have pre-Apache releases on github and our current effort mostly focused on a major refactoring of the codebase. Goal is to have it ready soon so that we don't need to support 2 distinct concurrent versions.

Signed off by mentor: phunt

Tashi
2011-October Tashi Incubator Status Report

Tashi has been incubating since September 2008.

The Tashi project aims to build a software infrastructure for cloud computing
on massive internet-scale datasets (what we call Big Data). The idea is to
build a cluster management system that enables the Big Data that are stored in
a cluster/data center to be accessed, shared, manipulated, and computed on by
remote users in a convenient, efficient, and safe manner.

Tashi has previously encompassed just the tools to manage virtual
machines using Xen and KVM, but is gaining the facility to hand out
physical machines as well.

Development activities have included:-
        * Zoni has been merged into the mainline code trunk
        * Additional capability added to Zoni
        * Implement hint to influence scheduler packing policy
        * Reject incorrect arguments to tashi-client to prevent
          unintended defaults from being used
        * Migrated to rpyc version 3.1
        * Add "free capacity" info function to Tashi
        * Support for auto creation of zoni tftp boot menus
        * Fixed deadlocks in clustermanager
        * Rewrite CM to concentrate decay handlers into one spot
        * Use Linux LVM for local scratch space creation
        * VMM is now authoritative to what is running
        * Retry deploying held VMs at a later time

The project is still working toward building a larger user and development
community. User groups have been identified in Ireland, Slovenia and Korea,
Malaysia, as well as at Georgia Tech.

Items to be resolved before graduation:
        * A stable branch exists which could be a release candidate, but
          the codebase is large and test hardware is currently in
          short supply. We are confident that the code in the stablefix
          branch will work if running QEMU emulation, Pickle or sqlite
          data storage, primitive scheduler. Xen, other data stores and
          schedulers have not been tested recently.
        * Should have example accounting code (data is kept, but
          intepetation is currently manual)

* Develop community diversity (Committers currently at Telefonica,
            Google and CMU)
Signed off by mentor:

VXQuery

The VXQuery Project implements a standard compliant XML Query processor. It has been in incubation since 2009-07-06.

Recent activities:

    * activity picked-up a little in August and September
    * considering direction of the project

Top issues before graduation:

    * Create a release
    * Build community (developer and users)

Signed off by mentor:

Wave

(project add text here)

Signed off by mentor: