

Apache Software Foundation

Digitaler Brutkasten

Keine Büros, keine Gebäude: Die Apache Software Foundation ist nahezu unsichtbar. Doch brütet sie Hunderte von Software-Großprojekten aus, bringt Tausende Entwickler zusammen. Und all das, ohne im Chaos zu enden und auf Innovation zu verzichten. Wie macht sie das? Ein Streifzug. *Kristian Kießling*



spiel benannt hat oder nach der nord-amerikanischen Stammesgruppe.

Steiler Aufstieg

Im April 1995 veröffentlichte die selbsternannte Apache Group, die aus acht Entwicklern bestand, eine erste Version 0.6.2 des Servers, im Dezember 1995 erschien die Version 1.0. Im Laufe eines Jahres überflügelten die Apache-Nutzerzahlen bereits die des NCSA HTTPd. Im Zuge dieser Entwicklung gründeten die Apache 1999 die Apache Software Foundation (ASF). Da waren bereits weitere Softwareprojekte im Apache-Umfeld entstanden, die sich etwa um die Perl- oder PHP-Module kümmerten.

Heute verwaltet die ASF 164 Toplevel-Projekte, mit 110 Subprojekten. Daneben gibt es 35 als Podlings bezeichnete Nachwuchsprojekte im Apache Incubator [8] sowie 39 Initiativen im Apache Lab [9]. Insgesamt sind das 348 Open-Source-Initiativen. Monatlich kommen durchschnittlich zwei neue Projekte hinzu (Abbildung 1), insgesamt zählt die ASF 4700 Codelieferanten [10].

Mitte der 90er entstand auch die Apache License 1.0, die damals noch verlangte, die Apache Group in Werbematerialien zu nennen. Der Passus fiel später weg. Die aktuelle Version 2.0 ist zur GPLv3 kompatibel, wenn ein Projekt Apache-2.0-Code in sein GPLv3-Projekt integriert. Umgekehrt funktioniert es nicht. Mehr zu dieser Problematik liefert [11].

Auf die Bühne

Die Apache Software Foundation bezeichnet ihre Organisationsform als Meritokratie. Grob vereinfacht heißt dies:

Apache-Projekte wie Hadoop [1], Cassandra [2], Tomcat [3] und Spark [4] erfreuen sich in den Rechenzentren seit Jahren größter Beliebtheit. Den Grundstein der Apache Software Foundation (ASF, [5]), deren Projekte der Schwerpunkt in diesem Linux-Magazin beleuchtet, legte aber ein schlanker, kleiner

HTTP-Server [6]. Ihn kann man im Softwarebereich wohl als äußerst populären Dauerbrenner bezeichnen (siehe Artikel im Schwerpunkt).

Vorläufer des Apache-Servers war der HTTPd der NCSA [7]. Wem die Abkürzung nichts sagt: Vom National Center for Supercomputing Applications stammt der wohl bekannteste der frühen Browser: Mosaic. Rob McCool arbeitete Anfang der 90er in der Institution und erhielt den Auftrag, einen Webserver zu entwickeln, der nicht so komplex ist wie der damals bevorzugt eingesetzte vom CERN. Den Code suchte er in dem noch jungen Internet zusammen.

Als McCool das NCSA verließ, um einen neuen Job anzutreten, wurde der Server ziemlich vernachlässigt. Lediglich ein paar Nutzer schickten noch eigene Patches, um ihn am Leben zu erhalten. Brian Behlendorf sammelt diese und so entstand „A patchy Server“. Unklar bleibt, ob sich Apache nach diesem Wort-

Inhalt

- 18 **Die Apache Software Foundation**
Vom Webserverprojekt zur millionenschweren Open-Source-Stiftung.
- 22 **Hadoop**
Big Data erklärt: Fischen im Datenmeer mit Hadoop 2 und Apache Spark.
- 26 **Apache Aurora**
Wer Dienste in Mesos terminieren möchte, greift zum Aurora-Framework.
- 30 **Jmeter**
Stressen, aber richtig! Jmeter ist Apaches Lasttest-Werkzeug.
- 34 **Apaches Webserver tunen**
Aus dem Webserver Nr. 1 lässt sich eine bessere Performance herauskitzeln.

Wer viel tut, hat auch viel zu sagen. Bestehende Mitglieder stimmen über die Aufnahme neuer Mitglieder ab, die Details regelt eine Satzung [12].

Die Apache-Projekte durchlaufen mehrere vorbestimmte Stadien (Abbildung 2). Wer sein Projekt in die Obhut der ASF übergeben hat, landet damit zunächst im Apache Incubator (Brutkasten, [8]). Stellt sich heraus, dass der Code darin aufblüht und sich eine Community um das Projekt bildet, wird es zum Toplevel-Projekt [13] und steht gleichwertig neben den anfangs genannten.

Um Innovationen anzukurbeln, bietet die ASF ihren Mitgliedern zugleich die Möglichkeit, mit Code zu experimentieren. Die dazugehörige Infrastruktur stellen die Apache Labs [9] bereit. Ist eines der Experimente erfolgreich, landet es im Incubator und folgt dem weiter oben vorgezeichneten Weg. Selbst an das Aufräumen denkt die Foundation: Nicht mehr betreute oder aktuelle Projekte landen auf dem Attic (Dachboden, [14]). Der Code bleibt so zugänglich und lässt sich wiederbeleben, falls sich Interessenten finden.

Wer sich aber für seinen nicht mehr betreuten Code ein zweites Leben im Apache-Projekt erhofft, missversteht das System. Die ASF akzeptiert Code nur, wenn dessen Besitzer einige Regeln beachtet und bestimmte Voraussetzungen erfüllt. Zu diesen gehören:

- Es muss sich um lauffähigen Code handeln. Das begründet die ASF damit, dass sich eine Community erst dann bilde, wenn der Code funktioniert. Stehe am Anfang zudem eine langwierige Designphase, würde es häufig zur Fragmentierung der Community kommen.
- Der Codebesitzer muss das Copyright der Software und das darin enthaltene geistige Eigentum im Rahmen eines Software Grant Agreement [15] der Apache-Stiftung übergeben. Nur dann können die User sicher sein, nicht irgendwann in einem Lock-in zu landen.
- Nicht zuletzt braucht das Projekt einen Sponsor innerhalb der ASF oder einen Officer. Diese Person agiert als Mentor, hilft dem Projekt und hält Kontakt zum Project Management Committee (PMC) des Projekts.

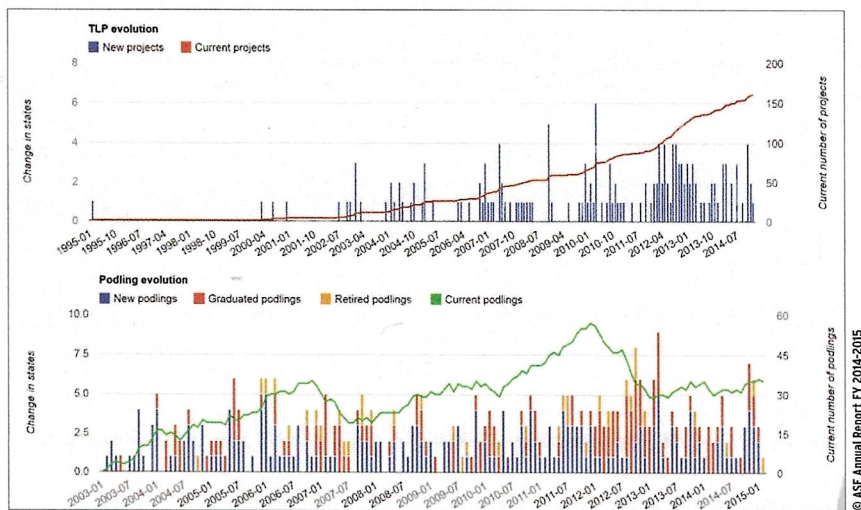


Abbildung 1: Sowohl bei den Toplevel-Projekten als auch im Incubator zeigt der Trend nach oben. Die Schwankungen im Incubator ergeben sich, weil reife Projekte in die Toplevel-Liga wechseln.

Die ASF glaubt an Hüte, also an Verantwortliche, die „den Hut aufhaben“: Für jedes Projekt gibt es nicht nur ein PMC, sondern auch ein Mitglied der Foundation, das als offizieller Ansprechpartner für das Projekt auftritt [16].

Gut versorgt

An der Spitze der ASF gibt es einen Vorstand mit neun Positionen. Sechzehn sogenannte Officer kümmern sich zudem um organisatorische Dinge, etwa das Marketing oder Fundraising. Für Letzteres gibt es sogar zwei Officers. Das ist kein Wunder, verfügt die Foundation doch über zahlreiche Sponsoren [17], die nicht nur die Infrastruktur, sondern auch die Arbeit der Foundation sowie Veranstaltungen, darunter die ASF-Konferenz Apachecon, finanzieren.

Sponsoren können zwischen unterschiedlichen Leveln wählen. Zu den Platinumsponsoren, die jeweils sportliche 100 000 US-Dollar im Jahr an die Foundation zahlen, gehören Google, Microsoft, Facebook, Citrix und Yahoo. Weitere bekannte Firmen treten als Gold-, Silber- oder Bronze-Sponsoren auf.

Das ist verständlich, weil die ASF zahlreiche Open-Source-Projekte verwaltet, die von diesen Firmen stammen und die dort täglich im Einsatz sind. Als Sponsoren versprechen sie sich eine Einflussnahme auf die Projektentwicklung und einen besseren Zugang zu Entwicklern. Wohin ihr Geld fließt, zeigen regelmäßig die Jahresfinanzberichte, der für 2014 etwa

unter [10]. Auf Konferenzen treffen die Firmenmitarbeiter auf freie Apache-Entwickler und es werden neue Entwicklungen vorgestellt und diskutiert.

Blick ins Labor

Solche Innovationen sollen unter anderem in den Apache Labs entstehen. Alle Lab-Projekte samt einer Kurzerklärung und einem Link auf den Code in Apaches-SVN-Repository liefert [18]. Die relevanten Informationen stecken entweder in einer »README.txt«-Datei, häufiger aber in einer XML-Datei namens »doap.rdf«. Das Linux-Magazin wirft im Folgenden einen Blick ins Labor und stellt drei der Projekte vor.

Das in Java, Python und Scala geschriebene Framework Alike [19] soll es erlauben, ähnliche Bilder automatisch zu erkennen. Dazu verbindet es mehrere Tools miteinander: Open CV für maschinelles Sehen, die Maschinen-Lernsoftware Apache Mahout sowie die Suchmaschine Lucene. Open CV soll so genannte visuelle Deskriptoren auf Bildern entdecken. Mahout zerlegt diese in visuelle Wörter, die Lucene dann indexiert.

Mouse [20] ist die Python-Implementierung des Apache Release Audit Tool (Rat, [21]). Konkret soll es Lizenzen im Quellcode eines Projekts erkennen, um Compliance zu gewährleisten. Anders als Rat besitzt Mouse dank Python aber keine Java-Abhängigkeiten.

Clouds [22] soll hingegen ein „Architektur-, Test- und Dokumentationsprojekt“

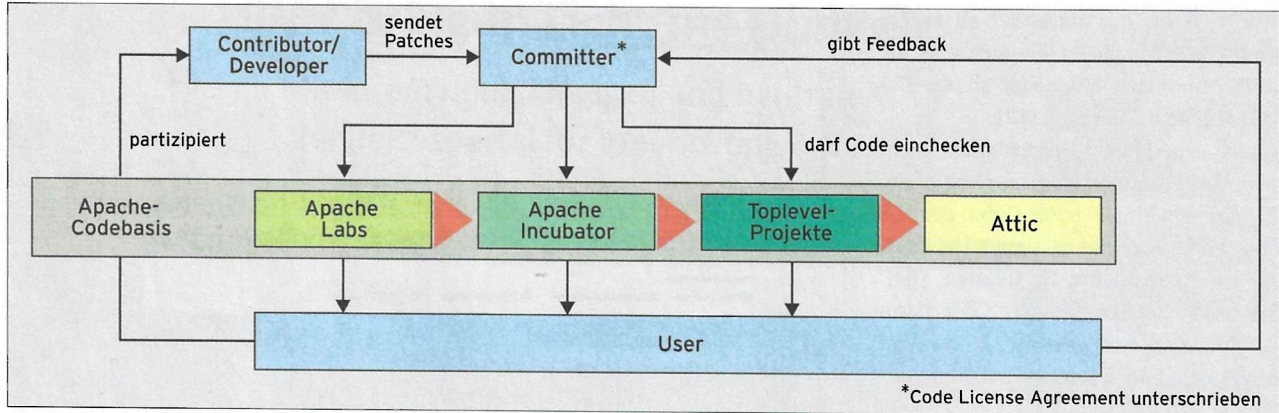


Abbildung 2: Firmen wie Google, Facebook und Twitter gehören häufig sowohl zu den Usern als auch zu den Entwicklern im Entwicklungsmodell der ASF.

sein, um Apache-Anwendungen in die Cloud zu bringen. Genauer soll es als Klebstoff zwischen Apaches Cloudstack und den Apache-Anwendungen dienen, um diese besser in den Cloudstack zu integrieren.

Beim Blick auf die weiteren Projekte zeigte sich aber auch, dass viele das Lab entweder schon wieder verlassen haben oder sie sich im Status »idle« befinden. Da diese Projekte zum Teil seit acht oder neun Jahren existieren, sind sie vermutlich schlicht außer Dienst.

Aus dem Brutkasten

Vielversprechender sehen dagegen die Projekte im Apache Incubator [23] aus, die den Sprung zum Toplevel-Projekt schaffen möchten. Zu den bekannteren im Brutkasten gehört Groovy [24]. Die objektorientierte Sprache für die Java-Plattform gibt es bereits seit 2003, aktuell ist die Version 2.4. Im Frühjahr zog Groovy als Inkubator-Projekt ins Lager der Apachen, nachdem die ehemaligen Wirte von Codehaus [25] das Geschäft einstellten.

Auch von Apache Wave [26] hat der eine oder andere womöglich schon gehört. Es handelt sich dabei um ein von Google angeschobenes Software-Framework für die Echtzeitkommunikation und Kollaboration über ein XMPP-basiertes Protokoll. Google trennte sich 2010 von dem Projekt, Apache akzeptierte Wave im Inkubator. Das Hauptprodukt ist das Server-basierte „Wave in a Box“ (WIAB). Es hostet und verteilt die von den Usern über eine Wave-Anwendung verfassten Waves. Dabei handelt es sich beispielsweise um Foren- oder E-Mail-Threads,

Instant-Messaging-Konversationen oder Wiki-Artikel.

Ripple [27] will sich beim Entwerfen von HTML-5-basierten Mobilanwendungen als Test- und Debug-Tool nützlich machen, unter anderem über einen Browser-basierten Mobiltelefon-Emulator. Aktuell unterstützt es Laufzeitumgebungen wie Cordova und Webworks.

Fazit

Während in Apaches Lab offenbar einige Karteileichen und nicht weiter verfolgte Experimente schlummern, scheinen die Projekte im Inkubator und die Toplevel-Projekte gut in Schuss. Auch die Apache Software Foundation als Ganzes erfreut sich offenbar guter Gesundheit und scheint finanziell und organisatorisch gangbare Wege gefunden zu haben, mit der wachsenden Zahl an Projekten umzugehen.

Die Organisationsform erinnert nicht zufällig an die in der Kernelentwicklung. Wie die Linux Foundation verdankt auch die ASF ihren Erfolg zunehmend dem finanziellen Engagement großer Firmen, die ihren Code dafür in die Stiftung auslagern. Für die Projekte erweist sich das offenbar als nützlich: Hauseigene Statistiken verweisen auf eine steigende Anzahl an Codelieferanten. ■

Infos

- [1] Hadoop: [http://hadoop.apache.org]
- [2] Cassandra: [http://cassandra.apache.org]
- [3] Tomcat: [http://tomcat.apache.org]
- [4] Spark: [http://spark.apache.org]
- [5] Apache Software Foundation: [http://apache.org]

- [6] Apache-Server: [http://httpd.apache.org]
- [7] Geschichte der ASF: [http://www.apache.org/history/timeline.html]
- [8] Incubator: [http://incubator.apache.org]
- [9] Labs: [http://labs.apache.org]
- [10] ASF-Fiskal-Jahresbericht 2014/2015: [https://s3.amazonaws.com/files-dist/AnnualReports/ASFAnnualReport-FY2014-2015FINAL.pdf]
- [11] Kompatibilität der Apache-Lizenz: [http://www.apache.org/licenses/GPL-compatibility.html]
- [12] Satzung der ASF: [http://www.apache.org/foundation/bylaws.html]
- [13] Toplevel-Projekte: [http://apache.org/#projects-list]
- [14] Attic: [http://attic.apache.org]
- [15] Software Grant Agreement: [http://www.apache.org/licenses/software-grant.txt]
- [16] ASF-Verantwortliche: [http://www.apache.org/foundation/]
- [17] Sponsoring: [http://www.apache.org/foundation/sponsorship.html]
- [18] Liste der Labs-Projekte: [http://labs.apache.org/labs.html]
- [19] Alike: [http://svn.apache.org/repos/asf/labs/alike/]
- [20] Mouse: [http://svn.apache.org/repos/asf/labs/mouse/]
- [21] Apache Rat: [http://creadur.apache.org/rat/]
- [22] Clouds: [http://svn.apache.org/repos/asf/labs/clouds]
- [23] Inkubator-Projekte: [http://incubator.apache.org/projects/index.html]
- [24] Groovy: [http://incubator.apache.org/projects/groovy.html]
- [25] Codehaus: [http://www.codehaus.org]
- [26] Apache Wave: [http://incubator.apache.org/projects/wave.html]
- [27] Ripple: [http://incubator.apache.org/projects/ripple.html]