

# An Examination of the Initiation, Organization, Participation, Leadership, and Control of Successful Open Source Software Development Projects

Michael P. Conlon  
Computer Science Department  
Slippery Rock University of Pennsylvania

# Overview

- Open Source/Free Software
- The Right to Fork
- Six Open Source/Free Software Projects
  - The Linux Kernel
  - The Apache Web Server
  - MySQL DBMS
  - PostgreSQL O/R DBMS
  - OpenOffice.org Suite
  - Alice
- Analysis

# Open Source/Free Software

- Open Source: source code is available.
- Free Software: Freedom to
  - Use the software for any purpose.
  - Study the software and see how it works.
  - Modify the software and contribute your improvements to the project and community.
  - Distribute the software to others.

(Free Software Foundation)
- Free software predates proprietary!

# Forking

- Definition: take the source code and run with it. Code diverges from code of the established leadership. Often there is a name change.
- Forks occur when team members are dissatisfied with leadership of their open source project.
- Usually the community follows one set of leaders. The other project withers.
- The threat of a fork protects the community from arbitrary decisions of development team leadership.
- E.g.: X.org vs. X-Free86

# Relevance of Forking

- Every open-source project lives with the risk of a fork.
- Project strategy must be developed with this in mind.
  - Project leadership must serve their community's needs effectively to avoid a fork.
  - Ability to fork makes lock-in impossible.
- Note how each following project tries to avoid forking.

# Examination of Six Open-Source Projects

- Linux, an operating-system kernel
- Apache, a Web server
- MySQL, a relational DBMS
- PostgreSQL, an object/relational DBMS
- OpenOffice.org, an office suite: word processor, spreadsheet, presentation package, desktop database, etc.
- Alice, a graphical programming environment for introducing beginners to computer programming

# Linux



- Initiated by Linus Torvalds, then a Finnish undergraduate student.
- He was soon joined by many hackers over the Internet.
- Eventually, major corporations (Digital, IBM, et al.) committed programmers to the project.
- Linux is licensed under the GPL.

# Linux Development Activity

Version 2.6.0:

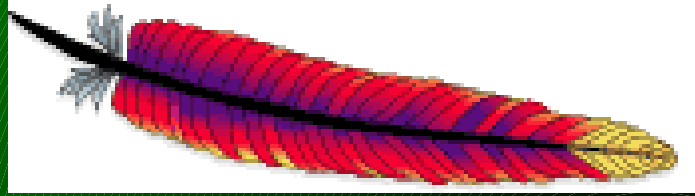
- 680 days development
- 27,149 changes
- Average 1.66 changes per hour
- 916 developers

(Kroah-Hartman, 2004)



# Linux Development Style

- Benevolent dictator
- Small cadre of lieutenants in charge of subsystems.
- No master plan: participants choose for themselves what to work on.
- No corporate ownership, in spite of much corporate involvement.
- Code releases “early and often.”



# Apache

- When NCSA dropped development and support of their Web server, *httpd*, some of the abandoned Webmasters started maintaining it themselves, under the project name, *Apache*.
- Soon Apache attracted many other Webmasters who contributed bug fixes and features.
- Today's Apache has no *httpd* code.
- Licensed under Apache license.

# Apache Development Community

- Development managed by Apache Software Foundation.
- Management by Project Management Committee and by consensus.
- Contributors graduate from user to developer to committer to PMC member.
- Promotion is based on competence.
- Claimed 800 committers.

# MySQL



- Started as a part of a data warehousing system in 1995.
- Has improved vastly since then: now competes with the big guys (Oracle, DB2, etc.)
- Adopted by many webmasters seeking an open-source database.

# MySQL Licensing

- Property of MySQL AB, which owns the source code and the MySQL trademark.
- Available under both open-source and proprietary licenses.
- Contributors must sign their code over to MySQL AB.
- MySQL AB makes its money on support for and licensing of its proprietary version.
- A fork would devastate this business model, so MySQL AB must listen to its customers.

# MySQL Development Team

- Anyone who signs a Contributor License Agreement can contribute.
- The number of contributors is not readily available, but they claim 11 million installations.
- Contributors are rewarded with licenses to the proprietary-licensed version and public recognition.
- Apparently there is no formal organization of contributors outside the company itself.

# PostgreSQL

- First open-source object-relational DBMS.
- BSD-licensed. This license permits commercialization.
- PostgreSQL had attracted a development community by 1995.
- Several commercial efforts have resulted in contributed improvements to the open-source version.

# PostgreSQL Community

- Maintained by the PostgreSQL Global Development Group.
- Code contributors are actively encouraged.
- Claimed 49 current and 17 former contributors.
- Submitted code is reviewed by experienced PostgreSQL developers.
- Presumably, one is invited to be a reviewer after establishing a reputation as a contributor.





# OpenOffice.org

- Began life as Star Office, a proprietary product.
- Purchased by Sun Microsystems in 1999, who stripped others' code out and made the rest open source, under the LGPL.
- This made it the first open-source, cross-platform, complete office suite, with an excellent ability to read and write MS Office files.

# Open Office.org Community

- Claimed 151 developers with write-access to the source code, with many others with lesser degrees of access.
- Organized similarly to Apache, with members, contributors, developers, and project leads.
- CollabNet hosts and helps manage the project, while Sun supports it.
- Contributors must give Sun joint-ownership of their code.

# Alice



- Initiated as a virtual reality development environment.
- Now, it's a highly-motivational system to teach computer programming to today's youth.
- Alice license: an open source license.

# Alice Community

- Users only.
- Development is in-house only. Outside contributions are not accepted.
- Bug reports are encouraged.
- Relatively slow development compared to the other projects: the last release was May, 2005.

# Forking Alice?

- Of all these projects, Alice seems most susceptible to a fork, because of its slow development pace.
- Reasons not to fork Alice:
  - Alice is great software!
  - The development team seems to have the best interest of its users at heart.
  - upcoming Alice 3.0 with Sims graphics. Sims license (???) may restrict forking.

# Summary

<b>System</b>	<b>No. of Developers</b>	<b>Age of Project (years)</b>	<b>Team Organization</b>	<b>Type of Organization Leading Project</b>	<b>Control</b>	<b>Open to New Developers?</b>	<b>Mature Technology?</b>
<b>Linux</b>	1722	16	Bazaar	none	Benevolent Dictator	Yes	Yes
<b>Apache</b>	800	13	Bazaar	Volunteer Organization	Committee Consensus	Yes	Yes
<b>MySQL</b>	1000(est.)	12	Bazaar	For-profit Company	Internal to MySQL AB	Yes	Yes
<b>PostgreSQL</b>	49	22	Bazaar	Volunteer Organization	Committee Consensus	Yes	Yes
<b>OpenOffice.org</b>	151	8	Bazaar	For-profit Company	Committee Consensus	Yes	Yes
<b>Alice</b>	10 (est.)	10	Cathedral	University Research Team	Internal to Stage3	No	No

# Analysis

- Each project started by single person or small co-located team.
- Most teams are large compared to equivalent proprietary-software teams.
- Most resulting software is stable and reliable, with rapid release of new versions and bug fixes.
- All projects seem to have the best interest of their users in mind.
- Alice is different!
  - Small team, with external contributions unwelcome
  - Some stability problems
  - Releases are infrequent
  - The only one of these projects creating state-of-the-art software.



# The End

# Thank you!