

CSE471 - Free and Open Source Software

Dayanand V

`v_dayanand@cb.amrita.edu`

Dept. of Computer Science and Engineering
Amrita School of Engineering
Amrita Vishwa Vidyapeetham

April 12, 2016

Credits

The credits for the contents of this presentation goes to Karl Fogel and his book “Producing Open Source Software”. You are as free to reuse and redistribute the contents of this presentation as I am in using Fogel’s book’s content.

Contents

- 1 Introduction
- 2 History
- 3 Overview of the development process
- 4 Technical Infrastructure
- 5 Packaging, Releasing, and Daily Development
- 6 Licenses, Copyrights and Patents
- 7 Introduction to GNU Autotools
- 8 Case Studies
 - The Apache Project
 - The Mozilla Project

Introduction

- About 90-95% free software projects fail. Why?
 - Unrealistic requirements
 - Vague specifications
 - Poor resource management
 - Insufficient design phases
- Aren't these the same reasons why closed-source projects fail?
 - Yes, and there's more.
- Sometimes, the developers do not appreciate the problems unique to open source development. Like?

Introduction

- More reasons why open source projects fail.
 - ① Expecting hordes of oompa-loompas to magically volunteer for you.
 - ② Expecting that releasing the source will be a cure of all its ills.
 - ③ Skimping on presentation and packaging to develop “more important stuff”.
 - ④ Expecting that the same management practices used for in-house development will work equally well on an open source project.
 - ⑤ Failures of “Cultural Navigation”.

“Open source does work, but it is most definitely not a panacea. If there’s a cautionary tale here, it is that you can’t take a dying project, sprinkle it with the magic pixie dust of “open source,” and have everything magically work out. Software is hard. The issues aren’t that simple.” -Jamie Zawinski, The Mozilla Project

Contents

- 1 Introduction
- 2 History**
- 3 Overview of the development process
- 4 Technical Infrastructure
- 5 Packaging, Releasing, and Daily Development
- 6 Licenses, Copyrights and Patents
- 7 Introduction to GNU Autotools
- 8 Case Studies
 - The Apache Project
 - The Mozilla Project

History

- At the dawn of commercial computers, ‘Software’ sounded closer to an accessory than a business asset.
- Customers (scientists, technicians etc) distributed patches to other customers and it was encouraged by the manufacturers.
- There was no hardware standardization and the manufacturer wanted machine-specific code and knowledge to spread as widely as possible.
- There was no Internet, and widespread frictionless sharing as we know it today, was not possible.

The Rise of Proprietary Software and Free Software

- The advent of better 'hardware' and 'high level' programming languages.
- Selling software began looking like a good strategy.
- If the user had freedom to modify the software and improve it, the 'added value' patches would be of lesser significance.
- Shared code can reach competitors.
- Ironically, the Internet was getting off the ground around the same time.

The Rise of Proprietary Software and Free Software - Conscious Resistance

- “Richard Stallman” happened.

“We did not call our software ”free software”, because that term did not yet exist; but that is what it was. Whenever people from another university or a company wanted to port and use a program, we gladly let them. If you saw someone using an unfamiliar and interesting program, you could always ask to see the source code, so that you could read it, change it, or cannibalize parts of it to make a new program.”
-Richard Stallman

- He went ahead and started the GNU Project and Free Software Foundation (FSF).
- GNU was aimed at developing a completely free and open computer operating system and body of application s/w.
- “The GNU General Public License (GPL) says that the code may be copied and modified without restriction and that both copies and derivative works must be distributed *under the same license as the original with no additional restrictions.*”

The Rise of Proprietary Software and Free Software - Conscious Resistance


- GNU slowly began gaining ground and was able to produce free replacements for many critical components of an OS.
- GNU Emacs came into being.
- The compiler collection GCC.
- By 1990s, GNU had produced most parts of a free OS except for the kernel.
- Linus Torvalds enters the scene with Linux, a kernel completed with the help of volunteers around the world.
- Becomes GNU Linux, the world's first free and open source operating system.

The Rise of Proprietary Software and Free Software - Accidental Resistance

- Berkeley Software Distribution (BSD) - University of California, Berkeley.
- Non ideological practice of free software development.
- X Window System developed at MIT allowed proprietary extensions over free software, but by itself was free.
- T_EX, developed by Donald Knuth offering free publishing quality typesetting system was released open source with just some naming restrictions on derivatives.

This presentation is built using ‘Beamer’ which is an extension of L^AT_EX, built on T_EX.

Free Versus Open Source

- Free? As in *free food*?
- Free as in *freedom*.
- Freedom to share and modify the source code for any purpose.
- Case study : Battle of browsers (1990s), Netscaps vs IE.
- Free software - Morality vs Profitability.
- Conflict of opinion.
- 1998 - the word “open source” is coined by Open Source Initiative (OSI ) to disambiguate “free” and aimed at giving ‘free software’ a good marketing.

The Situation Today

- Free software has become a culture of choice.
- So what persuades all these people to stick together long enough to produce something useful?
- The feeling that their connection to a project, and influence over it, is directly proportional to their contributions.
- Clearly, projects with corporate sponsorship and/or salaried developers need to be especially careful in this regard.

Contents

- 1 Introduction
- 2 History
- 3 Overview of the development process
- 4 Technical Infrastructure
- 5 Packaging, Releasing, and Daily Development
- 6 Licenses, Copyrights and Patents
- 7 Introduction to GNU Autotools
- 8 Case Studies
 - The Apache Project
 - The Mozilla Project

Look Around

“Every good work of software starts by scratching a developer’s personal itch.”
-Eric Raymond, ‘The Cathedral and the Bazaar’.

- A good software results when the programmer has a personal interest in seeing the problem solved.
- Today, we also have the phenomenon of organizations-for-profit, corporations, governments, non-profits, etc starting large, centrally-conceived open source projects from scratch.
- Case Study : Kuali Foundation
- Identifying potential problems that are faced by a large community.
- So, first, *look around!* github.com, openhub.net, sourceforge.net, directory.fsf.org

Start from What You Have

- Understanding clearly what the project is and what its not.
- Need for Documentation : Founders Vs New Comers.
- Investments : *hacktivation energy*.

Start from What You Have

- Choose a Good Name.
 - Gives an idea about what the project does.
 - Easy to remember.
 - Good manners, good legal sense.
 - Getting top-level domains with forwarding to a central home site.
 - Handle availability in Twitter, FB.

Start from What You Have

- Clear mission statement.
- State that the Project is 'Free'.
- A brief list of features and requirements.

Start from What You Have

- Development Status
 - Should reflect reality.
 - Alpha, Beta releases.
 - Availability for download.
 - Version Control and Bug Tracker Access.
 - GitHub.com - based on Git.
 - Rate of bug filing.
- Communication Channels
 - Mailing List.
 - Chat room.
 - IRC Channel.
- Developer Guidelines.

Start from What You Have

- Documentation.
 - Minimal criteria:
 - Should tell the reader clearly how much technical expertise they're expected to have.
 - Tells clearly how to setup, and make sure that they've installed the software correctly.
 - Give one tutorial-style example of how to do a common task.
 - Label the areas where the documentation is known to be incomplete.
 - Developer documentation.
- Demos, Screenshots, Videos, and Example Output.
- Hosting.

Choosing a License and Applying It

- The “Do-Anything” licenses (MIT License).
 - Asserts nominal copyright (without actually restricting copying).
 - Code comes with no warranty.
- The GNU General Public License (GPL).
 - Translates loosely to - “If you use my code, make sure you let others use yours.”

Choosing a License and Applying It

- Applying a license - Steps :
 - State the license clearly on the project's front page.
 - Put the full license text in a file called COPYING (or LICENSE) included with the source code.
 - Put a short notice in a comment at the top of each source file, naming :
 - Copyright date
 - Holder
 - The kind of license
 - Where to find the full text of the license.

Setting the Tone

- Avoid private discussions ($1 \ll n$).
- Zero tolerance to rudeness.
- Practice conspicuous code review.
- Be open from day one!
- Opening a formerly closed project.
- Announcing.

Contents

- 1 Introduction
- 2 History
- 3 Overview of the development process
- 4 Technical Infrastructure**
- 5 Packaging, Releasing, and Daily Development
- 6 Licenses, Copyrights and Patents
- 7 Introduction to GNU Autotools
- 8 Case Studies
 - The Apache Project
 - The Mozilla Project

Overview

- Website
- Mailing lists / Message forums
- Version control
- Bug tracking
- Real-time chat

Version Control and Website

- Canned Hosting Services.
- Github.com
- Setting up a Github account.
 - ① Go to <https://github.com/>
 - ② Sign up - Get yourself an account.
 - ③ Go to <https://guides.github.com/activities/hello-world/>
 - ④ Learn how to :
 - ① Create a 'repository'.
 - ② Open an 'issue'.
 - ③ Create a 'branch'.
 - ④ Make a 'commit'.
 - ⑤ Open a 'pull request'.
 - ⑥ 'Merge' a pull request.
 - ⑤ Go to <https://pages.github.com/> and follow the steps given there to setup a website for your project.

Mailing list, Chat and Wiki

- Google groups. *<https://www.groups.google.com>*
- Internet Relay Channel (IRC).
<http://www.irchelp.org/irchelp/altircfaq.html>
- Mediawiki. *<https://www.mediawiki.org>*

Contents

- 1 Introduction
- 2 History
- 3 Overview of the development process
- 4 Technical Infrastructure
- 5 Packaging, Releasing, and Daily Development**
- 6 Licenses, Copyrights and Patents
- 7 Introduction to GNU Autotools
- 8 Case Studies
 - The Apache Project
 - The Mozilla Project

Introduction

- Highway Repair analogy.
- Mutually independent tasks are parallelized.
- Release works can happen in parallel to development.

Releasing

- When does a release happen?
 - Some old bugs have been fixed and new ones added.
 - New features added.
 - New configuration options might have been added, changing the installation procedure.
 - Incompatible changes might have been introduced.

Releasing Numbering

- Unambiguously communicates the ordering of the release.
- Indicates the degree and nature of the changes in the release.
- Be consistent in choosing a numbering scheme, and stick with it.
- Some examples would be:
 - “Project 1.2.3”
 - “Project 1.2.3 (Alpha)” with the qualifier.
 - “Project 1.2.3 (Alpha 2)” with meta-qualifier.
- Qualifiers : Alpha, Beta, Stable, Unstable, Development, and RC (for “Release Candidate”) etc.
- “RC” always includes a numeric meta-qualifier, “Project 1.2.3 (RC 1)” is possible, whereas “Project 1.2.3 (RC)” is not.

Releasing Numbering

- $\langle \text{Project Name} \rangle \langle \text{Major Numbering} \rangle . \langle \text{Minor Numbering} \rangle . \langle \text{Micro Numbering} \rangle$
- “Project 2.10.17” is the eighteenth micro release (or patch release) in the eleventh minor release line within the second major release series.

The Simple Strategy

- Adopted by *APR project (Apache)*.
- Changes to the micro number only must be both forward- and backward-compatible.
 - Only bug fixes, or very small enhancements to existing features.
 - No new features introduced in a micro release.
- Changes to the minor number must be backward-compatible, but not necessarily forward-compatible.
 - New features can be introduced, but not too many.
- Changes to the major number mark compatibility boundaries. A new major release can be forward- and backward-incompatible.
 - Entire new feature sets can be expected.

The Even/Odd Strategy

- Parity of the minor number defines stability of the software.
- Even means stable, Odd means unstable.
- Increments in the micro number still indicate bug fixes (no new features).
- Increments in the major number still indicate big changes, new feature sets, etc.
- Compatibility guidelines are still preserved.

Release Branches

- The entire development tree will not be clean and ready for release at any time.
- A release branch is just a branch in the version control system (see branch), on which the code destined for this release can be isolated from mainline development.
- When a release branch is ‘stabilized’, it is time to tag a snapshot (a micro release) from the branch.
- ‘Stabilization’ is the process of deciding which changes will be in the release, which will not, and shaping the branch content accordingly.
- The ‘Release Owner’ sets the release rules and has the final say in the case of any conflict.
- Voting on changes.
- ‘Release Manager’.

Packaging

- Free software is distributed as source code or pre-built binary packages.
- Shipping formats:
 - *nix OSes - TAR compressed by /emphcompress, gzip, bzip, or bzip2.
 - MS Windows - /emphzip format.
- Name of the package contains the software's name plus the release number, plus the format suffixes appropriate for the archive type.

Layout

- Once uncompressed, the root directory should contain:
 - A 'README' plain text file describing :
 - what the software does.
 - what release this is.
 - link to the project's website etc.
 - An 'INSTALL' file giving instructions on how to build and install the software for all supporting operating systems.
 - A COPYING or LICENSE file, giving the software's terms of distribution.
 - A CHANGES file (sometimes called NEWS), explaining what's new in this release.

Compilation and Installation

- For compiled languages, in *NIX systems, the installation procedure of tarred source code typically would be like:
 - \$./configure (auto-detects the environment and prepares the build process)
 - \$ make (builds the software)
 - \$ make install (installs the software on the system)
- Installation of binary packages:
 - RPM system for RedHat GNU/Linux distros.
 - APT system for Debian GNU/Linux distros.

Testing

- Regression testing + Manual Testing.
- Once all check-list criteria are met, the packages are digitally signed using *GnuPG* or *PGP* etc.
- Once approved, the release should be placed into the project's download area, accompanied by the digital signatures, and by MD5/SHA1 checksums.
- The purpose of all this signing and checksumming is to give users a way to verify that the copy they receive has not been maliciously tampered with.

Candidate Releases

- Remember the qualifier 'RC'?
- Candidates releases are put out first before official release of a new version.
- The aim is to subject the code to wide testing before blessing it as an official release.
- If problems are found, they are fixed on the release branch and a new candidate release is rolled out.
- If there are no unacceptable bugs left, the last candidate release becomes the official release.

Announcing Releases

- The announcement should include:
 - The URL to the downloadable release tarball should always be accompanied with the MD5/SHA1 checksums and pointers to the digital signatures file.
 - Announcing in multiple forums/ mailing lists.
 - Reassurance on security.
 - The relevant portion of the CHANGES file.
 - Credits where it is due.

Maintaining Multiple Release Lines

- Most mature projects maintain multiple release lines in parallel.
- After 1.0.0 comes out, that line should continue with micro (bugfix) releases 1.0.1, 1.0.2, etc., until the project explicitly decides to end the line.
- Announcing the ‘end of life’ of release.
- Some projects set a window of time during which they pledge to support the previous release line.
- Some projects waits for the bug reports percentage to drops below a certain point to stop support.
- For each release, make sure to have a target version or target milestone available in the bug tracker, so people filing bugs will be able to do so against the proper release.

Security Releases

- A security release is a release made **SOLELY** to close a security vulnerability.
- This is friendly to the administrators who may need to deploy the security fix, but whose upgrade policy stipulates that they not deploy any other changes at the same time.
- An extra component can be added to an existing release to indicate that it contains only security changes (Eg: 1.1.2.1).

Releases and Daily Development

- Release Stabilization.
- Have each commit to be a single logical change, and don't mix unrelated changes in the same commit.
- If a change is too big or too disruptive to do in one commit, break it across N commits.
- Each commit is a well-partitioned subset of the overall change, and includes nothing unrelated to the overall change.

Planning Releases

- Deadlines, Schedules.
- Discussions, Compromises.
- Depending on the complexity of the release process, and the nature of the project, somewhere between every three and six months is usually about the right gap between releases.
- Maintenance lines may put out micro releases a bit faster, if there is demand for them.

Contents

- 1 Introduction
- 2 History
- 3 Overview of the development process
- 4 Technical Infrastructure
- 5 Packaging, Releasing, and Daily Development
- 6 Licenses, Copyrights and Patents**
- 7 Introduction to GNU Autotools
- 8 Case Studies
 - The Apache Project
 - The Mozilla Project

Aspects of Licenses

- Differences between licenses boil down to a few often recurring issues :
 - Compatibility with proprietary licenses.
 - Compatibility with other types of free licenses.
 - Enforcement of crediting.
 - Protection of trademark.
 - Patent snapback.
 - Protection of “artistic integrity”.

GPL and License Compatibility

- GPL Requirements:
 - Any derivative work must itself be distributed under the GPL.
 - No additional restrictions may be placed on the redistribution of either the original work or a derivative work.
- Incompatibilities:
 - Credit enforcement as a new policy cannot be applied on a GPL licensed code.
- Free Software Foundation maintains a *list* showing which licenses are compatible with the GPL and which are not.

Choosing a License

- Use one of the widely-used, well-recognized existing licenses.
- Advantages:
 - People won't feel they have to read the legalese in order to use your code, because they will have already done so for that license a long time ago.
 - They are the products of much thought and experience and the modern versions represent a great deal of accumulated legal and technical wisdom.

Choosing a License

- A list of licenses that meet the criteria: ¹
 - *GNU General Public License version 3 (GPL-3.0)*
 - *GNU Affero General Public License version 3 (AGPL-3.0)*
 - *Mozilla Public License 2.0 (MPL-2.0)*
 - *GNU Library or “Lesser” General Public License version 3 (LGPL-3.0)*
 - *Eclipse Public License 1.0 (EPL-1.0)*
 - *MIT license (MIT)*
 - *Apache License 2.0 (Apache-2.0)*
 - *BSD 2-Clause (“Simplified” or “FreeBSD”) license*

¹ **Assignment #3 : Study in detail and submit a report about the differences between the above licenses (with a clear tabular column) and where they would deem best to be used.**

Contributor Agreements

- Handling copyright ownership for free code and documentation that were contributed to by many people:
 - ➊ Ignore the issue of copyright entirely.
 - ➋ Collect a contributor license agreement (CLA) from each person who works on the project, explicitly granting the project the right to use that person's contributions.
 - ➌ Get actual copyright assignment (CA from contributors, so that the project (i.e., some legal entity, usually a nonprofit) is the copyright owner for everything.
- Even under centralized copyright ownership, the code remains free, because open source licenses do not give the copyright holder the right to retroactively proprietize all copies of the code.

Developer Certificates of Origin (DCO)

- An attestation that the contributor intends to contribute the enclosed code under the project's license, and that the contributor has the right to do so.
- The contributor indicates her understanding of the DCO once, early on, for example by emailing its text from her usual contribution address to a special archive at the project.
- The contributor includes a “Signed-Off-By:” line in her patches or commits, using the same identity, to indicate that the corresponding contribution is certified under the DCO.
- The DCO relies on the project's native open source license for any trademark or patent provisions, which in most cases is fine.

Proprietary Re-licensing

- An open source version of the software is available under the usual open source terms, while a proprietary version is available for a fee.
- Two types:
 - ① Selling exceptions to copyleft requirements.
 - Sells users a promise to not enforce the redistribution requirements of the open source version's license.
 - Eg:- MySQL Database Engine. GPL v2.
 - ② The freemium or open core model.
 - Uses an open source version to drive sales of a presumably fancier proprietary version.
- The catch here is that GPL's terms are something the copyright holder imposes on everyone else; the owner is therefore free to decide not to apply those terms to itself.

Proprietary Re-licensing

- Problems with proprietary re-licensing:
 - ① Discourages the normal dynamics of open source projects.
 - ② Any code contributors from outside the company are now effectively contributing to two distinct entities, one which they usually do not prefer.
 - ③ there is a deep motivational problem for open source projects that operate in the shadow of a proprietary re-licensed version.
 - ④ Hybrid “shakedown” models: anyone who makes commercially significant use of the code ends up being pressured to purchase a proprietary license as a way of protecting their commercial revenue stream from harassment.

- A trademarked name or symbol is a way for an entity (who owns or controls that trademark), to signal, in an easily recognizable way, that they approve of a particular product.
- Trademarks do not restrict copying, modification, or redistribution.
- Trademark is unrelated to copyright, and does not govern the same actions that copyright governs.
- Case study: Mozilla Firefox, the Debian Project, and Iceweasel.
- Case study: The GNOME Logo and the Fish Pedicure Shop.

Patents

- The only real threat against which the free software community cannot defend itself.
- Once someone has accused a free software project of infringing a patent, the project must either stop implementing that particular feature, or expose the project and its users to expensive and time-consuming lawsuits.
- Modern free software licenses generally have clauses to combat, or at least mitigate, the dangers arising from software patents.
- Usually these clauses work by automatically revoking the overall open source license for any party who makes a patent infringement claim based on either the work as a whole, or based on the claimant's code contributions to the project.

Contents

- 1 Introduction
- 2 History
- 3 Overview of the development process
- 4 Technical Infrastructure
- 5 Packaging, Releasing, and Daily Development
- 6 Licenses, Copyrights and Patents
- 7 Introduction to GNU Autotools**
- 8 Case Studies
 - The Apache Project
 - The Mozilla Project

Introduction to GNU Autotools

Introduction to GNU Autotools

- Outcome of the demand for a de-facto standard for compiling and installing software.

Introduction to GNU Autotools

- Outcome of the demand for a de-facto standard for compiling and installing software.
- Autoconf, Automake, and Libtool were developed separately, to make tackling the problem of software configuration more manageable by partitioning it.

Introduction to GNU Autotools

- Outcome of the demand for a de-facto standard for compiling and installing software.
- Autoconf, Automake, and Libtool were developed separately, to make tackling the problem of software configuration more manageable by partitioning it.
- By-products of the GNU/Linux project.

GNU Autotools and what they do!

- Autoconf:

GNU Autotools and what they do!

- Autoconf:
 - A tool that makes your packages more portable by performing tests to discover system characteristics before the package is compiled.

GNU Autotools and what they do!

- Autoconf:
 - A tool that makes your packages more portable by performing tests to discover system characteristics before the package is compiled.
 - The source code adapts to these differences.

GNU Autotools and what they do!

- Autoconf:
 - A tool that makes your packages more portable by performing tests to discover system characteristics before the package is compiled.
 - The source code adapts to these differences.
- Automake:

GNU Autotools and what they do!

- Autoconf:
 - A tool that makes your packages more portable by performing tests to discover system characteristics before the package is compiled.
 - The source code adapts to these differences.
- Automake:
 - A tool for generating ‘Makefile’s—descriptions of what to build—that conform to a number of standards.

GNU Autotools and what they do!

- Autoconf:
 - A tool that makes your packages more portable by performing tests to discover system characteristics before the package is compiled.
 - The source code adapts to these differences.
- Automake:
 - A tool for generating ‘Makefile’s—descriptions of what to build—that conform to a number of standards.
 - It substantially simplifies the process of describing the organization of a package.

GNU Autotools and what they do!

- Autoconf:
 - A tool that makes your packages more portable by performing tests to discover system characteristics before the package is compiled.
 - The source code adapts to these differences.
- Automake:
 - A tool for generating ‘Makefile’s—descriptions of what to build—that conform to a number of standards.
 - It substantially simplifies the process of describing the organization of a package.
 - Performs additional functions such as dependency tracking between source files.

GNU Autotools and what they do!

- Autoconf:
 - A tool that makes your packages more portable by performing tests to discover system characteristics before the package is compiled.
 - The source code adapts to these differences.
- Automake:
 - A tool for generating ‘Makefile’s—descriptions of what to build—that conform to a number of standards.
 - It substantially simplifies the process of describing the organization of a package.
 - Performs additional functions such as dependency tracking between source files.
- Libtool:

GNU Autotools and what they do!

- Autoconf:
 - A tool that makes your packages more portable by performing tests to discover system characteristics before the package is compiled.
 - The source code adapts to these differences.
- Automake:
 - A tool for generating ‘Makefile’s—descriptions of what to build—that conform to a number of standards.
 - It substantially simplifies the process of describing the organization of a package.
 - Performs additional functions such as dependency tracking between source files.
- Libtool:
 - A command line interface to the compiler and linker that makes it easy to portably generate static and shared libraries, regardless of the platform it is running on.

History of Autotools

- The advent and popularity of Unix.

History of Autotools

- The advent and popularity of Unix.
- POSIX standards.

History of Autotools

- The advent and popularity of Unix.
- POSIX standards.
- Need for an organized approach to handle the differences.

History of Autotools

- The advent and popularity of Unix.
- POSIX standards.
- Need for an organized approach to handle the differences.
- Earlier tools :

History of Autotools

- The advent and popularity of Unix.
- POSIX standards.
- Need for an organized approach to handle the differences.
- Earlier tools :
 - The Metaconfig program, by Larry Wall, Harlan Stenn, and Raphael Manfredi.
 - The Cygnus 'configure' script, by K. Richard Pixley, and the original GCC 'configure' script, by Richard Stallman.
 - The GNU Autoconf package, by David MacKenzie.
 - Imake, part of the X Window system.

History of Autotools

- The advent and popularity of Unix.
- POSIX standards.
- Need for an organized approach to handle the differences.
- Earlier tools :
 - The Metaconfig program, by Larry Wall, Harlan Stenn, and Raphael Manfredi.
 - The Cygnus ‘configure’ script, by K. Richard Pixley, and the original GCC ‘configure’ script, by Richard Stallman.
 - The GNU Autoconf package, by David MacKenzie.
 - Imake, part of the X Window system.
- All these tools had a *configuration step* and a *build step*.

History of Autotools

- The advent and popularity of Unix.
- POSIX standards.
- Need for an organized approach to handle the differences.
- Earlier tools :
 - The Metaconfig program, by Larry Wall, Harlan Stenn, and Raphael Manfredi.
 - The Cygnus ‘configure’ script, by K. Richard Pixley, and the original GCC ‘configure’ script, by Richard Stallman.
 - The GNU Autoconf package, by David MacKenzie.
 - Imake, part of the X Window system.
- All these tools had a *configuration step* and a *build step*.
- The configuration step would generate ‘Makefile’s, and perhaps other files, which would then be used during the build step using the ‘make’ command.

History of Autotools

- The advent and popularity of Unix.
- POSIX standards.
- Need for an organized approach to handle the differences.
- Earlier tools :
 - The Metaconfig program, by Larry Wall, Harlan Stenn, and Raphael Manfredi.
 - The Cygnus ‘configure’ script, by K. Richard Pixley, and the original GCC ‘configure’ script, by Richard Stallman.
 - The GNU Autoconf package, by David MacKenzie.
 - Imake, part of the X Window system.
- All these tools had a *configuration step* and a *build step*.
- The configuration step would generate ‘Makefile’s, and perhaps other files, which would then be used during the build step using the ‘make’ command.
- The use of ‘feature tests’ made software compatible across OS variants without change in source code.

History of Autotools

- By 1994, Autoconf was a solid framework for handling the differences between Unix variants. But,

History of Autotools

- By 1994, Autoconf was a solid framework for handling the differences between Unix variants. But,
 - Program developers still had to write large 'Makefile.in' files in order to use it.

History of Autotools

- By 1994, Autoconf was a solid framework for handling the differences between Unix variants. But,
 - Program developers still had to write large ‘Makefile.in’ files in order to use it.
 - The ‘configure’ script generated by autoconf would transform the ‘Makefile.in’ file into a ‘Makefile’ used by the make program.

History of Autotools

- By 1994, Autoconf was a solid framework for handling the differences between Unix variants. But,
 - Program developers still had to write large ‘Makefile.in’ files in order to use it.
 - The ‘configure’ script generated by autoconf would transform the ‘Makefile.in’ file into a ‘Makefile’ used by the make program.

History of Autotools

- Since most programs are built in much the same way, there was a great deal of duplication in ‘Makefile.in’ files.

History of Autotools

- Since most programs are built in much the same way, there was a great deal of duplication in ‘Makefile.in’ files.
- Also, the GNU project developed a reasonably complex set of standards for ‘Makefile’s, and it was easy to get some of the details wrong.

History of Autotools

- Since most programs are built in much the same way, there was a great deal of duplication in ‘Makefile.in’ files.
- Also, the GNU project developed a reasonably complex set of standards for ‘Makefile’s, and it was easy to get some of the details wrong.
- These factors led to the development of Automake.

History of Autotools

- Since most programs are built in much the same way, there was a great deal of duplication in ‘Makefile.in’ files.
- Also, the GNU project developed a reasonably complex set of standards for ‘Makefile’s, and it was easy to get some of the details wrong.
- These factors led to the development of Automake.
 - The developer writes files named ‘Makefile.am’; these use a simpler syntax than ordinary ‘Makefile’s. automake reads the ‘Makefile.am’ files and produces ‘Makefile.in’ files.

History of Autotools

- Since most programs are built in much the same way, there was a great deal of duplication in ‘Makefile.in’ files.
- Also, the GNU project developed a reasonably complex set of standards for ‘Makefile’s, and it was easy to get some of the details wrong.
- These factors led to the development of Automake.
 - The developer writes files named ‘Makefile.am’; these use a simpler syntax than ordinary ‘Makefile’s. automake reads the ‘Makefile.am’ files and produces ‘Makefile.in’ files.
 - A script generated by autoconf converts these ‘Makefile.in’ files into ‘Makefile’s.

History of Autotools

- Since most programs are built in much the same way, there was a great deal of duplication in ‘Makefile.in’ files.
- Also, the GNU project developed a reasonably complex set of standards for ‘Makefile’s, and it was easy to get some of the details wrong.
- These factors led to the development of Automake.
 - The developer writes files named ‘Makefile.am’; these use a simpler syntax than ordinary ‘Makefile’s. automake reads the ‘Makefile.am’ files and produces ‘Makefile.in’ files.
 - A script generated by autoconf converts these ‘Makefile.in’ files into ‘Makefile’s.
 - automake also adds any rules required by the GNU ‘Makefile’ standards.

History of Autotools

- Since most programs are built in much the same way, there was a great deal of duplication in ‘Makefile.in’ files.
- Also, the GNU project developed a reasonably complex set of standards for ‘Makefile’s, and it was easy to get some of the details wrong.
- These factors led to the development of Automake.
 - The developer writes files named ‘Makefile.am’; these use a simpler syntax than ordinary ‘Makefile’s. automake reads the ‘Makefile.am’ files and produces ‘Makefile.in’ files.
 - A script generated by autoconf converts these ‘Makefile.in’ files into ‘Makefile’s.
 - automake also adds any rules required by the GNU ‘Makefile’ standards.

History of Autotools

- Unix added support for shared libraries.

History of Autotools

- Unix added support for shared libraries.
- Advantages of using shared libraries.

History of Autotools

- Unix added support for shared libraries.
- Advantages of using shared libraries.
- Development of Libtool began.

History of Autotools

- Unix added support for shared libraries.
- Advantages of using shared libraries.
- Development of Libtool began.
 - Libtool is a collection of shell scripts which handle the differences between shared library generation and use on different systems.
 - It is closely tied to Automake, although it is possible to use it independently.
- Over time, Libtool has been enhanced to support more Unix variants and to provide an interface for standardizing shared library features.

History of Autotools

- In 1995, Microsoft released Windows 95 which became widely popular.

History of Autotools

- In 1995, Microsoft released Windows 95 which became widely popular.
- Autoconf and Libtool provided a framework to support portability (from a single source code base) to Windows as well.

History of Autotools

- In 1995, Microsoft released Windows 95 which became widely popular.
- Autoconf and Libtool provided a framework to support portability (from a single source code base) to Windows as well.
- To support this, the GNU bash shell was ported to Windows as Cygwin which implements the basic Unix API in Windows.

History of Autotools

- In 1995, Microsoft released Windows 95 which became widely popular.
- Autoconf and Libtool provided a framework to support portability (from a single source code base) to Windows as well.
- To support this, the GNU bash shell was ported to Windows as Cygwin which implements the basic Unix API in Windows.
- Autoconf and Libtool support Windows directly, using either the Cygwin interface or the Visual C++ tools from Microsoft.

History of Autotools

- In 1995, Microsoft released Windows 95 which became widely popular.
- Autoconf and Libtool provided a framework to support portability (from a single source code base) to Windows as well.
- To support this, the GNU bash shell was ported to Windows as Cygwin which implements the basic Unix API in Windows.
- Autoconf and Libtool support Windows directly, using either the Cygwin interface or the Visual C++ tools from Microsoft.
- Automake has also been ported to Windows but it requires Perl to be installed.

Running ‘configure’

- Command line options.

- https://www.sourceware.org/autobook/autobook/autobook_14.html#Configuring

Running ‘configure’

- Command line options.
 - https://www.gnu.org/autobook/autobook/autobook_14.html#Configuring
 - ‘--cache-file=file’

Running ‘configure’

- Command line options.
 - https://www.gnu.org/autobook/autobook/autobook_14.html#Configuring
 - ‘--cache-file=file’
 - ‘--help’

Running ‘configure’

- Command line options.
 - https://www.gnu.org/autobook/autobook/autobook_14.html#Configuring
 - ‘--cache-file=file’
 - ‘--help’
 - ‘--quiet’ or ‘--silent’

Running ‘configure’

- Command line options.
 - https://www.gnu.org/autobook/autobook/autobook_14.html#Configuring
 - ‘--cache-file=file’
 - ‘--help’
 - ‘--quiet’ or ‘--silent’
 - ‘--disable-⟨feature⟩’

Running ‘configure’

- Command line options.
 - https://www.gnu.org/autobook/autobook/autobook_14.html#Configuring
 - ‘--cache-file=file’
 - ‘--help’
 - ‘--quiet’ or ‘--silent’
 - ‘--disable-⟨feature⟩’
 - ‘--enable-feature[=arg]’etc..
- Source tree, Build tree and Install tree.

Files generated by configure

- 'config.cache'

Files generated by configure

- 'config.cache'
- 'config.log'

Files generated by configure

- 'config.cache'
- 'config.log'
- 'config.status'

Files generated by configure

- 'config.cache'
- 'config.log'
- 'config.status'
- 'config.h'

Files generated by configure

- 'config.cache'
- 'config.log'
- 'config.status'
- 'config.h'
- 'Makefile'

Files generated by configure

- 'config.cache'
- 'config.log'
- 'config.status'
- 'config.h'
- 'Makefile'

Makefiles and targets

- Makefiles contain specifications of dependencies between files and how to resolve those dependencies
- Targets - the task to be performed by the Makefile:
 - make *all*
 - make *check*
 - make *install*
 - make *clean* etc.

Make files

- The ‘make’ program attempts to bring a target up to date by bring all of the target’s dependencies up to date (the dependent has a more recent timestamp than the target).
- Dependencies and dependency graph.

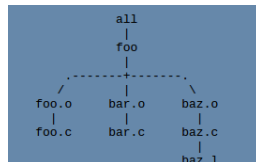
```
all: foo

foo: foo.o bar.o baz.o

.c.o:
    $(CC) $(CFLAGS) -c $< -o $@

.l.c:
    $(LEX) $< && mv lex.yy.c $@
```

(a) Makefile



(b) Dependency Graph

- When leaf nodes are found in the dependency graph, the ‘Makefile’ must include a set of shell commands to bring the dependent up to date with the dependency.

Makefiles

- Each of the shell commands are run in their own sub-shell and, unless the ‘Makefile’ instructs make otherwise.
- Each command must exit with an exit code of 0 to indicate success.
- Target rules can be written which are executed unconditionally (by specifying that the target has no dependents).

Writing Makefiles

- Syntax:

```
target1: dep1 dep2 ... depN
<tab>    cmd1
<tab>    cmd2
<tab>    ...
<tab>    cmdN

target2: dep4 dep5
<tab>    cmd1
<tab>    cmd2

dep4 dep5:
<tab>    cmd1
```

- Command prefixes - '@', '-'
- Macros
 - Starts with a dollar sign.
 - Eg:- \$CC
 - Define a make variable using a 'var=value' syntax: 'CC = ec++' (Default value of \$CC is 'cc')
 - '\$@' and '\$/' represent the names of the target and the first dependency for the rule in which they appear

Writing Makefiles

- Example:

```
all:    dummy
        @echo "$@" depends on dummy"

dummy:
        touch $@
```

(c) File

```
$ make
touch dummy
all depends on dummy
```

(d) Output

- Suffix Rules: wildcard pattern that can match targets
- Example:

```
.c.o:   $(CC) $(CFLAGS) -c $< -o $@
```

- This rule will match any target that ends in '.o' and are said to always be dependent on '.c'

Contents

- 1 Introduction
- 2 History
- 3 Overview of the development process
- 4 Technical Infrastructure
- 5 Packaging, Releasing, and Daily Development
- 6 Licenses, Copyrights and Patents
- 7 Introduction to GNU Autotools
- 8 Case Studies**
 - The Apache Project
 - The Mozilla Project

The credits for the contents of the following slides goes to Joseph Feller, Brian Fitzgerald, Scott A. Hissam, and Karim R. Lakhani and their book “Perspectives on Free and Open Source Software”. Most of the contents are taken from the book and I suggest that you read the book to get a detailed view.

Introduction to the development process

- Began in February 1995 as a combined effort to coordinate existing fixes to the NCSA httpd program developed by Rob McCool.

Introduction to the development process

- Began in February 1995 as a combined effort to coordinate existing fixes to the NCSA httpd program developed by Rob McCool.
- All existing features, and many new ones, were ported to the new architecture designed by Robert Thau.

Introduction to the development process

- Began in February 1995 as a combined effort to coordinate existing fixes to the NCSA httpd program developed by Rob McCool.
- All existing features, and many new ones, were ported to the new architecture designed by Robert Thau.
- Formal release of Apache httpd 1.0 happened in January 1996.

Introduction to the development process

- Began in February 1995 as a combined effort to coordinate existing fixes to the NCSA httpd program developed by Rob McCool.
- All existing features, and many new ones, were ported to the new architecture designed by Robert Thau.
- Formal release of Apache httpd 1.0 happened in January 1996.
- Apache began with a conscious attempt to solve the process issues first, before development even started.

Roles and Responsibilities

- The Apache Group (AG), the informal organization of people responsible for guiding the development of the Apache HTTP Server Project consisted entirely of volunteers.

Roles and Responsibilities

- The Apache Group (AG), the informal organization of people responsible for guiding the development of the Apache HTTP Server Project consisted entirely of volunteers.
- Required a development and decision-making process that emphasized decentralized workspaces and asynchronous communication.

Roles and Responsibilities

- The Apache Group (AG), the informal organization of people responsible for guiding the development of the Apache HTTP Server Project consisted entirely of volunteers.
- Required a development and decision-making process that emphasized decentralized workspaces and asynchronous communication.
- AG used e-mail lists exclusively to communicate with each other, and a minimal quorum voting system for resolving conflicts.

Roles and Responsibilities

- The Apache Group (AG), the informal organization of people responsible for guiding the development of the Apache HTTP Server Project consisted entirely of volunteers.
- Required a development and decision-making process that emphasized decentralized workspaces and asynchronous communication.
- AG used e-mail lists exclusively to communicate with each other, and a minimal quorum voting system for resolving conflicts.
- AG members are people who have contributed for an extended period of time, usually more than six months, and are nominated for core developer membership and then voted on by the existing members.

Roles and Responsibilities

- Core developers include the subset of AG that is active in development (usually 4 to 6 in any given week) and the developers who are on the cusp of being nominated to AG membership (usually 2 to 3).

Roles and Responsibilities

- Core developers include the subset of AG that is active in development (usually 4 to 6 in any given week) and the developers who are on the cusp of being nominated to AG membership (usually 2 to 3).
- Each AG member can vote on the inclusion of any code change, and has commit access to CVS.

Roles and Responsibilities

- Core developers include the subset of AG that is active in development (usually 4 to 6 in any given week) and the developers who are on the cusp of being nominated to AG membership (usually 2 to 3).
- Each AG member can vote on the inclusion of any code change, and has commit access to CVS.
- Votes are generally reserved for major changes that would affect other developers who are adding or changing functionality.

The Development Process

- Each Apache core developer *iterates* through a common series of actions while working on the software source.

The Development Process

- Each Apache core developer *iterates* through a common series of actions while working on the software source.
 - ① Discovering that a problem exists or new functionality is needed.

The Development Process

- Each Apache core developer *iterates* through a common series of actions while working on the software source.
 - ① Discovering that a problem exists or new functionality is needed.
 - ② Determining whether a volunteer will work on the issue.

The Development Process

- Each Apache core developer *iterates* through a common series of actions while working on the software source.
 - ➊ Discovering that a problem exists or new functionality is needed.
 - ➋ Determining whether a volunteer will work on the issue.
 - ➌ Identifying a solution.

The Development Process

- Each Apache core developer *iterates* through a common series of actions while working on the software source.
 - ➊ Discovering that a problem exists or new functionality is needed.
 - ➋ Determining whether a volunteer will work on the issue.
 - ➌ Identifying a solution.
 - ➍ Developing and testing the code within their local copy of the source.

The Development Process

- Each Apache core developer *iterates* through a common series of actions while working on the software source.
 - ① Discovering that a problem exists or new functionality is needed.
 - ② Determining whether a volunteer will work on the issue.
 - ③ Identifying a solution.
 - ④ Developing and testing the code within their local copy of the source.
 - ⑤ Presenting the code changes to the AG for review.

The Development Process

- Each Apache core developer *iterates* through a common series of actions while working on the software source.
 - ① Discovering that a problem exists or new functionality is needed.
 - ② Determining whether a volunteer will work on the issue.
 - ③ Identifying a solution.
 - ④ Developing and testing the code within their local copy of the source.
 - ⑤ Presenting the code changes to the AG for review.
 - ⑥ Committing the code and documentation to the repository.

The Development Process

- Each Apache core developer *iterates* through a common series of actions while working on the software source.
 - ➊ Discovering that a problem exists or new functionality is needed.
 - ➋ Determining whether a volunteer will work on the issue.
 - ➌ Identifying a solution.
 - ➍ Developing and testing the code within their local copy of the source.
 - ➎ Presenting the code changes to the AG for review.
 - ➏ Committing the code and documentation to the repository.
- It is generally preferred that the entire set of changes needed to solve a particular problem or add a particular enhancement be applied in a single commit.

Identifying Work to Be Done

- Avenues through which problems can be reported and enhancement proposed:

Identifying Work to Be Done

- Avenues through which problems can be reported and enhancement proposed:
 - Developer mailing list - where new features and patches for bugs are discussed.

Identifying Work to Be Done

- Avenues through which problems can be reported and enhancement proposed:
 - Developer mailing list - where new features and patches for bugs are discussed.
 - The bug reporting system (BUGDB) - where bugs are reported (usually with no patch).

Identifying Work to Be Done

- Avenues through which problems can be reported and enhancement proposed:
 - Developer mailing list - where new features and patches for bugs are discussed.
 - The bug reporting system (BUGDB) - where bugs are reported (usually with no patch).
 - Usenet newsgroups associated with the Apache products.

Identifying Work to Be Done

- Avenues through which problems can be reported and enhancement proposed:
 - Developer mailing list - where new features and patches for bugs are discussed.
 - The bug reporting system (BUGDB) - where bugs are reported (usually with no patch).
 - Usenet newsgroups associated with the Apache products.
- Change requests on the mailing list are given the highest priority because:

Identifying Work to Be Done

- Avenues through which problems can be reported and enhancement proposed:
 - Developer mailing list - where new features and patches for bugs are discussed.
 - The bug reporting system (BUGDB) - where bugs are reported (usually with no patch).
 - Usenet newsgroups associated with the Apache products.
- Change requests on the mailing list are given the highest priority because:
 - Most probably from a member of the development community.

Identifying Work to Be Done

- Avenues through which problems can be reported and enhancement proposed:
 - Developer mailing list - where new features and patches for bugs are discussed.
 - The bug reporting system (BUGDB) - where bugs are reported (usually with no patch).
 - Usenet newsgroups associated with the Apache products.
- Change requests on the mailing list are given the highest priority because:
 - Most probably from a member of the development community.
 - More likely to contain sufficient information to analyze the request or contain a patch to solve the problem.

Identifying Work to Be Done

- Avenues through which problems can be reported and enhancement proposed:
 - Developer mailing list - where new features and patches for bugs are discussed.
 - The bug reporting system (BUGDB) - where bugs are reported (usually with no patch).
 - Usenet newsgroups associated with the Apache products.
- Change requests on the mailing list are given the highest priority because:
 - Most probably from a member of the development community.
 - More likely to contain sufficient information to analyze the request or contain a patch to solve the problem.
- Problems such as compilation or build problems, are typically found first by core developers and either fixed immediately or reported and handled on the mailing list.

Identifying Work to Be Done

- In order to keep track of the project status, an agenda file (STATUS) is stored in each products repository, containing a list of high-priority problems, open issues among the developers, and release plans.

Identifying Work to Be Done

- In order to keep track of the project status, an agenda file (STATUS) is stored in each products repository, containing a list of high-priority problems, open issues among the developers, and release plans.
- Anyone with Web or e-mail access can enter and categorize requests in the BUGDB by severity and topic area.

Identifying Work to Be Done

- In order to keep track of the project status, an agenda file (STATUS) is stored in each products repository, containing a list of high-priority problems, open issues among the developers, and release plans.
- Anyone with Web or e-mail access can enter and categorize requests in the BUGDB by severity and topic area.
- Once entered, the request is posted to a separate mailing list and can be appended to via e-mail replies or edited directly by the core developers.

Identifying Work to Be Done

- In order to keep track of the project status, an agenda file (STATUS) is stored in each products repository, containing a list of high-priority problems, open issues among the developers, and release plans.
- Anyone with Web or e-mail access can enter and categorize requests in the BUGDB by severity and topic area.
- Once entered, the request is posted to a separate mailing list and can be appended to via e-mail replies or edited directly by the core developers.
- When a problem from any source is repaired, the BUGDB is searched for reports associated with that problem so that they can be included in the change report and closed.

Identifying Work to Be Done

- In general, the AG relies on interested volunteers and the community at large to recognize promising enhancements and real problems, and to take the time to report them to the BUGDB or forward them directly to the developer mailing list.

Identifying Work to Be Done

- In general, the AG relies on interested volunteers and the community at large to recognize promising enhancements and real problems, and to take the time to report them to the BUGDB or forward them directly to the developer mailing list.
- The problems reported on released versions of the server are recorded in BUGDB.

Identifying Work to Be Done

- In general, the AG relies on interested volunteers and the community at large to recognize promising enhancements and real problems, and to take the time to report them to the BUGDB or forward them directly to the developer mailing list.
- The problems reported on released versions of the server are recorded in BUGDB.
- In order for a proposed change actually to be made, an AG member must ultimately be persuaded it is needed or desirable.

Identifying Work to Be Done

- In general, the AG relies on interested volunteers and the community at large to recognize promising enhancements and real problems, and to take the time to report them to the BUGDB or forward them directly to the developer mailing list.
- The problems reported on released versions of the server are recorded in BUGDB.
- In order for a proposed change actually to be made, an AG member must ultimately be persuaded it is needed or desirable.
- “Showstoppers” problems that are sufficiently serious (in the view of a majority of AG members) that a release cannot go forward until they are solved are always addressed.

Assigning and Performing Development Work

- Generally, a problem/enhancement that has found favor in AG would be assigned to a volunteer who is ready to work on the problem.

Assigning and Performing Development Work

- Generally, a problem/enhancement that has found favor in AG would be assigned to a volunteer who is ready to work on the problem.
- This gives the core developers enough time to concentrate on their own modules which are part of the core functionality of the software.

Assigning and Performing Development Work

- Generally, a problem/enhancement that has found favor in AG would be assigned to a volunteer who is ready to work on the problem.
- This gives the core developers enough time to concentrate on their own modules which are part of the core functionality of the software.
- The core developers obtain an implicit “code ownership” of parts of the server that they are known to have created or to have maintained consistently.

Assigning and Performing Development Work

- Generally, a problem/enhancement that has found favor in AG would be assigned to a volunteer who is ready to work on the problem.
- This gives the core developers enough time to concentrate on their own modules which are part of the core functionality of the software.
- The core developers obtain an implicit “code ownership” of parts of the server that they are known to have created or to have maintained consistently.
- After work assignment, the next step is identifying a solution.

Assigning and Performing Development Work

- Generally, a problem/enhancement that has found favor in AG would be assigned to a volunteer who is ready to work on the problem.
- This gives the core developers enough time to concentrate on their own modules which are part of the core functionality of the software.
- The core developers obtain an implicit “code ownership” of parts of the server that they are known to have created or to have maintained consistently.
- After work assignment, the next step is identifying a solution.
- The different solutions identified are forwarded to the mailing list to get feedback before finalizing on one.

Prerelease Testing

- The local copy of the code is often tested on the developers own local server (similar to unit testing).

Prerelease Testing

- The local copy of the code is often tested on the developers own local server (similar to unit testing).
- The thoroughness of the test depends on the judgement and expertise of the developer.

Prerelease Testing

- The local copy of the code is often tested on the developers own local server (similar to unit testing).
- The thoroughness of the test depends on the judgement and expertise of the developer.
- Reviews are required before and after committing the change.

Inspections

- After unit testing, the core developer either commits the changes directly or produces a “patch” and posts it to the developer mailing list for review.

Inspections

- After unit testing, the core developer either commits the changes directly or produces a “patch” and posts it to the developer mailing list for review.
- In general, changes to a stable release require review before being committed, whereas changes to development releases are reviewed after the change is committed.

Inspections

- After unit testing, the core developer either commits the changes directly or produces a “patch” and posts it to the developer mailing list for review.
- In general, changes to a stable release require review before being committed, whereas changes to development releases are reviewed after the change is committed.
- If approved, the patch can be committed to the source by any of the developers, although in most cases it is preferred that the originator of the change also perform the commit.

Inspections

- After unit testing, the core developer either commits the changes directly or produces a “patch” and posts it to the developer mailing list for review.
- In general, changes to a stable release require review before being committed, whereas changes to development releases are reviewed after the change is committed.
- If approved, the patch can be committed to the source by any of the developers, although in most cases it is preferred that the originator of the change also perform the commit.
- All of the core developers are responsible for reviewing the apache-cvs mailing list to ensure that the changes are appropriate.

Managing Releases

- Release manager would be a core developer who is responsible for identifying the critical problems that prevent the release (if any) and determining when the software becomes stable.

Managing Releases

- Release manager would be a core developer who is responsible for identifying the critical problems that prevent the release (if any) and determining when the software becomes stable.
- Release manager also controls access to the repository.

Managing Releases

- Release manager would be a core developer who is responsible for identifying the critical problems that prevent the release (if any) and determining when the software becomes stable.
- Release manager also controls access to the repository.
- Enforces tying up all loose ends.

Managing Releases

- Release manager would be a core developer who is responsible for identifying the critical problems that prevent the release (if any) and determining when the software becomes stable.
- Release manager also controls access to the repository.
- Enforces tying up all loose ends.
- The role of release manager is rotated among the core developers with the most experience with the project.

Quantitative Results

- Although they cannot be fully generalized, the metrics considered are:

Quantitative Results

- Although they cannot be fully generalized, the metrics considered are:
 - Size of the community.

Quantitative Results

- Although they cannot be fully generalized, the metrics considered are:
 - Size of the community.
 - Work distribution.

Quantitative Results

- Although they cannot be fully generalized, the metrics considered are:
 - Size of the community.
 - Work distribution.
 - Problem reporting.

Quantitative Results

- Although they cannot be fully generalized, the metrics considered are:
 - Size of the community.
 - Work distribution.
 - Problem reporting.
 - Code ownership.

Quantitative Results

- Although they cannot be fully generalized, the metrics considered are:
 - Size of the community.
 - Work distribution.
 - Problem reporting.
 - Code ownership.
 - Defect density.

Quantitative Results

- Although they cannot be fully generalized, the metrics considered are:
 - Size of the community.
 - Work distribution.
 - Problem reporting.
 - Code ownership.
 - Defect density.
 - Time to resolve problem reports.

Quantitative Results - Size of the community

- The overall number of volunteers who actively contributed
- almost 400.

Quantitative Results - Size of the community

- The overall number of volunteers who actively contributed - almost 400.
- Around 3060 different people submitted 3,975 problem reports, whereas 458 individuals submitted 591 reports that subsequently caused a change to the Apache code or documentation.

Quantitative Results - Size of the community

- The overall number of volunteers who actively contributed - almost 400.
- Around 3060 different people submitted 3,975 problem reports, whereas 458 individuals submitted 591 reports that subsequently caused a change to the Apache code or documentation.
- There were 2654 individuals who submitted 3384 reports that could not be traced to a code change.

Quantitative Results - Size of the community

- The overall number of volunteers who actively contributed - almost 400.
- Around 3060 different people submitted 3,975 problem reports, whereas 458 individuals submitted 591 reports that subsequently caused a change to the Apache code or documentation.
- There were 2654 individuals who submitted 3384 reports that could not be traced to a code change.
- The number of contributors who submitted problem fixes - 182 (695 fixes).

Quantitative Results - Size of the community

- The overall number of volunteers who actively contributed - almost 400.
- Around 3060 different people submitted 3,975 problem reports, whereas 458 individuals submitted 591 reports that subsequently caused a change to the Apache code or documentation.
- There were 2654 individuals who submitted 3384 reports that could not be traced to a code change.
- The number of contributors who submitted problem fixes - 182 (695 fixes).
- The number of contributors who submitted code - 249 (6092 code submissions).

Quantitative Results - Size of the community

- The overall number of volunteers who actively contributed - almost 400.
- Around 3060 different people submitted 3,975 problem reports, whereas 458 individuals submitted 591 reports that subsequently caused a change to the Apache code or documentation.
- There were 2654 individuals who submitted 3384 reports that could not be traced to a code change.
- The number of contributors who submitted problem fixes - 182 (695 fixes).
- The number of contributors who submitted code - 249 (6092 code submissions).

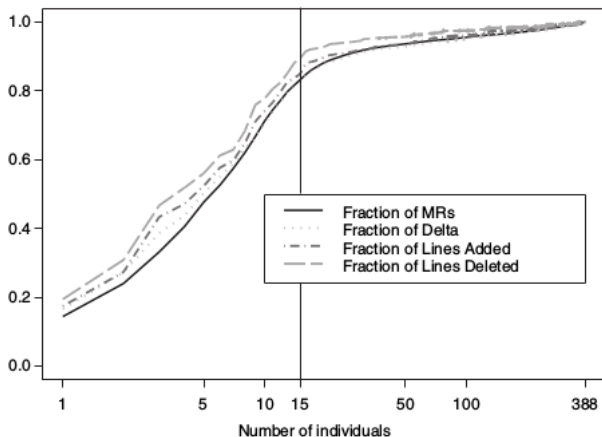
Quantitative Results - Work distribution

- Statistics on Apache and five commercial projects:

	MRs (K)	Delta (K)	Lines added (K)	Years	Developers
Apache	6	18	220	3	388
A	3.3	129	5,000	3	101
B	2.5	18	1,000	1.5	91
C	1.1	2.8	81	1.3	17
D	0.2	0.7	21	1.7	8
E	0.7	2.4	90	1.5	16

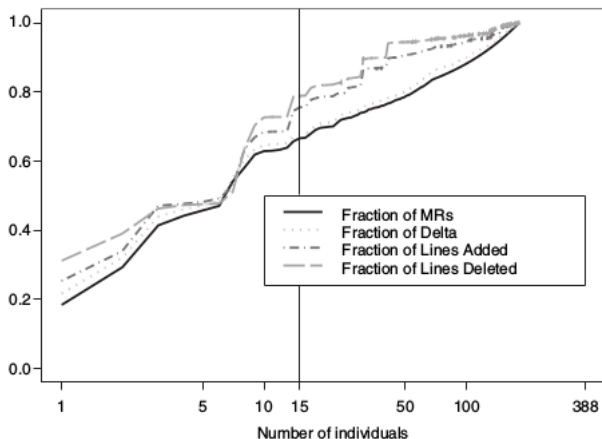
Quantitative Results - Work distribution

- Cumulative distribution of contributions to the code base:



Quantitative Results - Work distribution

- Cumulative distribution of contributions to bug fixes:



Quantitative Results - Problem Reporting

- Reported to BUGDB.

Quantitative Results - Problem Reporting

- Reported to BUGDB.
- Total number of distinct reports - 3975.

Quantitative Results - Problem Reporting

- Reported to BUGDB.
- Total number of distinct reports - 3975.
- The top 15 problems reporters submitted only 213 or 5% of PRs.

Quantitative Results - Problem Reporting

- Reported to BUGDB.
- Total number of distinct reports - 3975.
- The top 15 problems reporters submitted only 213 or 5% of PRs.
- Of the top 15 problem reporters only 3 are also core developers.

Quantitative Results - Problem Reporting

- Reported to BUGDB.
- Total number of distinct reports - 3975.
- The top 15 problems reporters submitted only 213 or 5% of PRs.
- Of the top 15 problem reporters only 3 are also core developers.
- The significant role of system tester is reserved almost exclusively for the wide community of Apache users.

Quantitative Results - Code Ownership

- Out of 42 .c files with more than 30 changes, 40 had at least 2 (and 20 had at least 4) developers making more than 10% of the changes.

Quantitative Results - Code Ownership

- Out of 42 .c files with more than 30 changes, 40 had at least 2 (and 20 had at least 4) developers making more than 10% of the changes.
- Rather than any single individual writing all the code for a given module, those in the core group have a sufficient level of mutual trust that they contribute code to various modules as needed.

Quantitative Results - Code Ownership

- Out of 42 .c files with more than 30 changes, 40 had at least 2 (and 20 had at least 4) developers making more than 10% of the changes.
- Rather than any single individual writing all the code for a given module, those in the core group have a sufficient level of mutual trust that they contribute code to various modules as needed.
- Code ownership is considered to be more of a matter of recognition than a strictly enforced ability to make commits to partitions of the code base.

Quantitative Results - Defect Density

- Measures - Postrelease defects per thousand lines of code added to the repo etc

Quantitative Results - Defect Density

- Measures - Postrelease defects per thousand lines of code added to the repo etc

Comparison of Defect Density Measures

Measure	Apache	A	C	D	E
Postrelease Defects/KLOCA	2.64	0.11	0.1	0.7	0.1
Postrelease Defects/KDelta	40.8	4.3	14	28	10
Postfeature test Defects/KLOCA	2.64	*	5.7	6.0	6.9
Postfeature test Defects/KDelta	40.8	*	164	196	256

- Although the user-perceived defect density of the Apache product is inferior to that of the commercial products, the defect density of the code before system test is much lower.

Quantitative Results - Defect Density

- Measures - Postrelease defects per thousand lines of code added to the repo etc

Comparison of Defect Density Measures

Measure	Apache	A	C	D	E
Postrelease Defects/KLOCA	2.64	0.11	0.1	0.7	0.1
Postrelease Defects/KDelta	40.8	4.3	14	28	10
Postfeature test Defects/KLOCA	2.64	*	5.7	6.0	6.9
Postfeature test Defects/KDelta	40.8	*	164	196	256

- Although the user-perceived defect density of the Apache product is inferior to that of the commercial products, the defect density of the code before system test is much lower.
- Indicates that fewer defects are injected into the code, or that other defect-finding activities such as inspections are conducted more frequently or more effectively

Quantitative Results - Time to Resolve Problem Reports

- The distribution of Apache PR resolution interval is approximated by its empirical distribution function that maps the interval in days to proportion of PRs resolved within that interval.

Quantitative Results - Time to Resolve Problem Reports

- The distribution of Apache PR resolution interval is approximated by its empirical distribution function that maps the interval in days to proportion of PRs resolved within that interval.
- Fifty percent of PRs are resolved within a day, 75 percent within 42 days, and 90 percent within 140 days.

Quantitative Results - Time to Resolve Problem Reports

- The distribution of Apache PR resolution interval is approximated by its empirical distribution function that maps the interval in days to proportion of PRs resolved within that interval.
- Fifty percent of PRs are resolved within a day, 75 percent within 42 days, and 90 percent within 140 days.
- Problems are resolved depending on priority, time period, and whether the PR causes a change to the code.

Quantitative Results - Time to Resolve Problem Reports

- The distribution of Apache PR resolution interval is approximated by its empirical distribution function that maps the interval in days to proportion of PRs resolved within that interval.
- Fifty percent of PRs are resolved within a day, 75 percent within 42 days, and 90 percent within 140 days.
- Problems are resolved depending on priority, time period, and whether the PR causes a change to the code.

Introduction to the development process

- Initially, the public expectations out of Mozilla project was set high and the number of volunteers were not meeting the expectations.

Introduction to the development process

- Initially, the public expectations out of Mozilla project was set high and the number of volunteers were not meeting the expectations.
- The first one year was cumbersome and demotivating for an impatient volunteer since there was no working product/prototype ready and there was a lack of adequate support from Netscape.

Introduction to the development process

- Initially, the public expectations out of Mozilla project was set high and the number of volunteers were not meeting the expectations.
- The first one year was cumbersome and demotivating for an impatient volunteer since there was no working product/prototype ready and there was a lack of adequate support from Netscape.
- After the documentation was improved, tutorials were written, and the development tools and processes refined, participation started slowly to increase.

Roles and Responsibilities

- Mozilla.org staff coordinate and guide the project, provide process and a few of them engage in coding.

Roles and Responsibilities

- Mozilla.org staff coordinate and guide the project, provide process and a few of them engage in coding.
- There are roles like community QA, milestone releases, Website tools and maintenance, tools like Bugzilla that assist developers.

Roles and Responsibilities

- Mozilla.org staff coordinate and guide the project, provide process and a few of them engage in coding.
- There are roles like community QA, milestone releases, Website tools and maintenance, tools like Bugzilla that assist developers.
- Decision-making authority for various modules is delegated to individuals in the development community who are close to that particular code.

Roles and Responsibilities

- Mozilla.org staff coordinate and guide the project, provide process and a few of them engage in coding.
- There are roles like community QA, milestone releases, Website tools and maintenance, tools like Bugzilla that assist developers.
- Decision-making authority for various modules is delegated to individuals in the development community who are close to that particular code.
- Directories and files within a particular module can be added or changed by getting the permission of the module owner.

Roles and Responsibilities

- Mozilla.org staff coordinate and guide the project, provide process and a few of them engage in coding.
- There are roles like community QA, milestone releases, Website tools and maintenance, tools like Bugzilla that assist developers.
- Decision-making authority for various modules is delegated to individuals in the development community who are close to that particular code.
- Directories and files within a particular module can be added or changed by getting the permission of the module owner.
- mozilla.org has the ultimate decision-making authority, and retains the right to designate and remove module owners, and to resolve all conflicts that arise.

Identifying Work to Be Done

- A roadmap document is maintained, that specifies what will be included in future releases, as well as dates for which releases are scheduled.

Identifying Work to Be Done

- A roadmap document is maintained, that specifies what will be included in future releases, as well as dates for which releases are scheduled.
- Enough steps are taken to ensure that development community is able to comment and participate in decisions.

Identifying Work to Be Done

- A roadmap document is maintained, that specifies what will be included in future releases, as well as dates for which releases are scheduled.
- Enough steps are taken to ensure that development community is able to comment and participate in decisions.
- Anyone can report bugs or request enhancements using the Bugzilla problem-reporting tool, thereby duplicate bug reports are mostly avoided.

Assigning and Performing Development Work

- Experts who write code focus on their own areas.

Assigning and Performing Development Work

- Experts who write code focus on their own areas.
- Volunteers search for Bugzilla to identify bugs or enhancements.

Assigning and Performing Development Work

- Experts who write code focus on their own areas.
- Volunteers search for Bugzilla to identify bugs or enhancements.
- Fixes are often submitted as attachments to Bugzilla problem reports.

Assigning and Performing Development Work

- Experts who write code focus on their own areas.
- Volunteers search for Bugzilla to identify bugs or enhancements.
- Fixes are often submitted as attachments to Bugzilla problem reports.
- Developers can mark Bugzilla items with a “helpwanted” keyword if they think an item is worth doing but dont themselves have the resources or all the required expertise.

Assigning and Performing Development Work

- Experts who write code focus on their own areas.
- Volunteers search for Bugzilla to identify bugs or enhancements.
- Fixes are often submitted as attachments to Bugzilla problem reports.
- Developers can mark Bugzilla items with a “helpwanted” keyword if they think an item is worth doing but don't themselves have the resources or all the required expertise.
- Discussions can also be found in Mozilla news groups, which may give development community members ideas about where to contribute.

Prerelease Testing

- Runs a daily minimal “smoke test” on the build for several major platforms, in order to ensure the build is sufficiently stable to allow development work on it to proceed.

Prerelease Testing

- Runs a daily minimal “smoke test” on the build for several major platforms, in order to ensure the build is sufficiently stable to allow development work on it to proceed.
- If the smoke test identifies bugs, they are posted daily so that developers are aware of any serious problems in the build.

Inspections

- Two stages :
 - 1 Module owners review a patch in the context of the module.

Inspections

- Two stages :
 - ① Module owners review a patch in the context of the module.
 - ② A smaller designated group (referred to as superreviewers, who are technically highly accomplished) review a patch for its interaction with the code base as a whole before it is checked in.

Managing Releases

- Mozilla runs a continuous build process (Tinderbox) that shows what parts of the code have issues for certain builds and under certain platform in which changes and corresponding authors are highlighted.

Managing Releases

- Mozilla runs a continuous build process (Tinderbox) that shows what parts of the code have issues for certain builds and under certain platform in which changes and corresponding authors are highlighted.
- It also produces binaries nightly and issues 'Milestones' approximately monthly.

Managing Releases

- Mozilla runs a continuous build process (Tinderbox) that shows what parts of the code have issues for certain builds and under certain platform in which changes and corresponding authors are highlighted.
- It also produces binaries nightly and issues ‘Milestones’ approximately monthly.
- These Milestone decisions are made by a designated group, known as “drivers@mozilla.org”, with input from the community.

Quantitative Results - Size of the Community

- 486 people who contributed code and 412 who contributed code to PR fixes that were incorporated.

Quantitative Results - Size of the Community

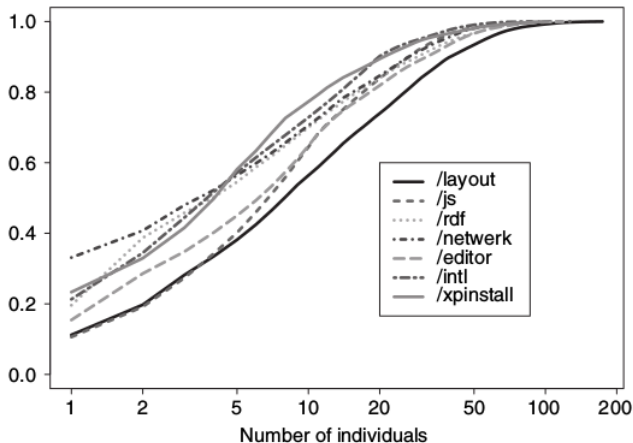
- 486 people who contributed code and 412 who contributed code to PR fixes that were incorporated.

Sizes of Apache, five commercial projects, and seven Mozilla modules

	MRs (K)	Delta (K)	Lines added (K)	Years	Developers
Apache	6	18	220	3	388
A	3.3	129	5,000	3	101
B	2.5	18	1,000	1.5	91
C	1.1	2.8	81	1.3	17
D	0.2	0.7	21	1.7	8
E	0.7	2.4	90	1.5	16
/layout	12.7	42	800	2.6	174
/js	4.6	14	308	2.6	127
/rdf	4.1	12	274	2	123
/network	3.2	10	221	1.6	106
/editor	2.9	8	203	2	118
/intl	2	5	118	1.8	87
/xpininstall	1.9	5	113	1.7	102

Quantitative Results - Size of the Community

- The cumulative contributions made to five different Mozilla modules.



Quantitative Results - Code Ownership

- Unlike Apache, code ownership is enforced.

Quantitative Results - Code Ownership

- Unlike Apache, code ownership is enforced.
- The module owner is responsible for fielding bug reports, enhancement requests, and patch submissions in order to facilitate good development.

Quantitative Results - Code Ownership

- Unlike Apache, code ownership is enforced.
- The module owner is responsible for fielding bug reports, enhancement requests, and patch submissions in order to facilitate good development.
- MRs are to be reviewed by the module owner.

Quantitative Results - Code Ownership

- Unlike Apache, code ownership is enforced.
- The module owner is responsible for fielding bug reports, enhancement requests, and patch submissions in order to facilitate good development.
- MRs are to be reviewed by the module owner.
- To manage check in privileges, Mozilla uses a Web-based tool called despot.

Quantitative Results - Defects

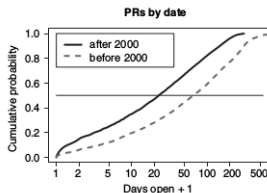
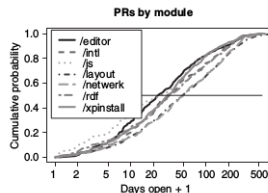
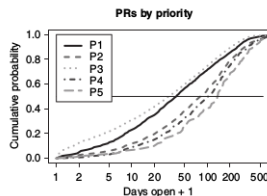
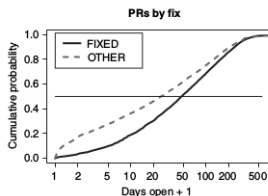
- Comparison of post-feature-test defect density measures.

Comparison of post-feature-test defect density measures

Module	#PR/KDelta	#PR/KLOC added
Apache	40.8	2.6
C	164	5.7
D	196	6.0
E	256	6.9
/layout	51	2.8
/js	19	0.7
/rdf	27	1.4
/network	42	3.1
/editor	44	2.5
/intl	20	1.6
/xpinstall	56	4.0

Quantitative Results - Time to Resolve Problem Reports

- Plots the cumulative distribution of the interval for all resolved PRs.



Comparison - Apache vs Mozilla

Hypotheses proposed for comparison:

- ① Open source developments will have a core of developers who control the code base. This core will be no larger than 10 to 15 people, and will create approximately 80 percent or more of the new functionality.
- ② For projects that are so large that 10 to 15 developers cannot write 80 percent of the code in a reasonable time frame, a strict code ownership policy will have to be adopted to separate the work of additional groups, creating, in effect, several related OSS projects.

Comparison - Apache vs Mozilla

Hypotheses proposed for comparison:

- ③ In successful open source developments, a group larger by an order of magnitude than the core will repair defects, and a yet larger group (by another order of magnitude) will report problems.
- ④ Open source developments that have a strong core of developers but never achieve large numbers of contributors beyond that core will be able to create new functionality, but will fail because of a lack of resources devoted to finding and repairing defects.

Comparison - Apache vs Mozilla

Hypotheses proposed for comparison:

- ⑤ Defect density in open source releases will generally be lower than commercial code that has only been feature-tested; that is, received a comparable level of testing.
- ⑥ In successful open source developments, the developers will also be users of the software.
- ⑦ OSS developments exhibit very rapid responses to customer problems.

Assignment

Assignment #4 : Explore how the seven hypothesis proposed by the authors for comparison are justified or not justified by the Apache and Mozilla projects and report YOUR OWN conclusions based on the observations.

Deadline - 12th April, 8.45 AM