



# **Bsc Thesis - On the reasons for founding OSS projects**

Wolfgang Sprenger

Institut für Informatik

FU Berlin

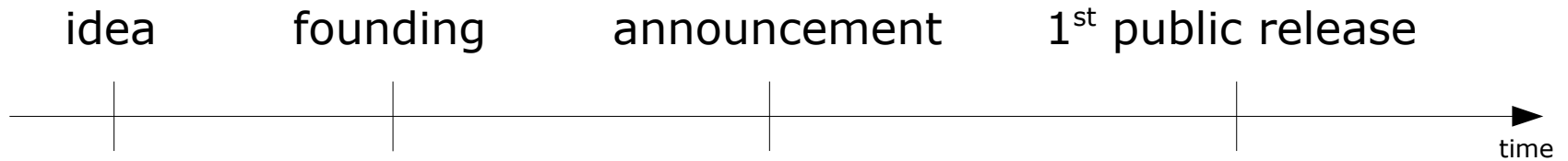
05.02.2009

- I. Motivation for this study
- II. What is founding?
- III. Methodology Part I (Literature review)
- IV. Methodology Part II (E-mail Study)
- V. Result
- VI. Analysis
- VII. Conclusion
- VIII. Discussion

- FLOSS, WIDI and other studies show only why people participate in Open Source Software projects\*
- No other study looks at the reasons for founding
- Is 'Scratching an Itch' the only reason (Raymond)?

\* OSS project = OSI license, open ended, ownership

## II. What is founding?



- Differentiate between 'idea', 'founding', 'announcement' and '1<sup>st</sup> public release'
- Special case: out-sourcing fork
- Founding in this study: 1<sup>st</sup> public release

!Caution: This will exclude all projects which never have a release.

- Literature review:
  - Academic literature
  - 10 well known OSS projects and their founding stories
  - Reasons for founding derived from community literature

## III.a) 10 well known OSS projects

- Apache Webserver
- Mozilla (Firefox)
- Gcc
- Samba
- Cups
- Linux Kernel
- Eclipse
- Open Office
- LaTeX (TeX)
- BSD

- Netscape lost more and more market share to IE
- Netscape Communications made two important announcements on January 23rd, 1998:
  - First, that the Netscape Communicator product would be available free of charge
  - Second, that the source code for Communicator would also be free
- Mozilla was the original code name

(Source: <http://www.mozilla.org/mission.html>)

## III.c) Reasons for founding:

- Scratching an itch
- Plan to close existing software gap
- Strategic reasons
- Exercise
- Academic research project
- Fundamental reason
- Existing project growing too big
- To reach a goal fast



## III.d) 10 well known OSS projects

- Apache Webserver -> Scratching an itch
- Mozilla (Firefox) -> Strategic reasons
- Gcc -> Scratching an itch
- Samba -> Scratching an itch
- Cups -> Scratching an itch
- Linux Kernel -> Exercise
- Eclipse -> Strategic reasons
- Open Office -> Strategic reasons
- LaTeX (TeX) -> Scratching an itch
- BSD -> Fundamental reason

- E-mail Study (100 each):
  - Debian packet list
  - Sourceforge.net list (through sql queries)
  - Intersection of Debian paket list and Sourceforge.net list
- What data is interessting:
  - Reasons for founding
  - Date of 'having the idea', founding and 1<sup>st</sup> release
  - Number of developers
  - What kind of communication tools are used

## IV.b) Debian packet list

- approx. 20.000 Pakete
- list is downloadable from website

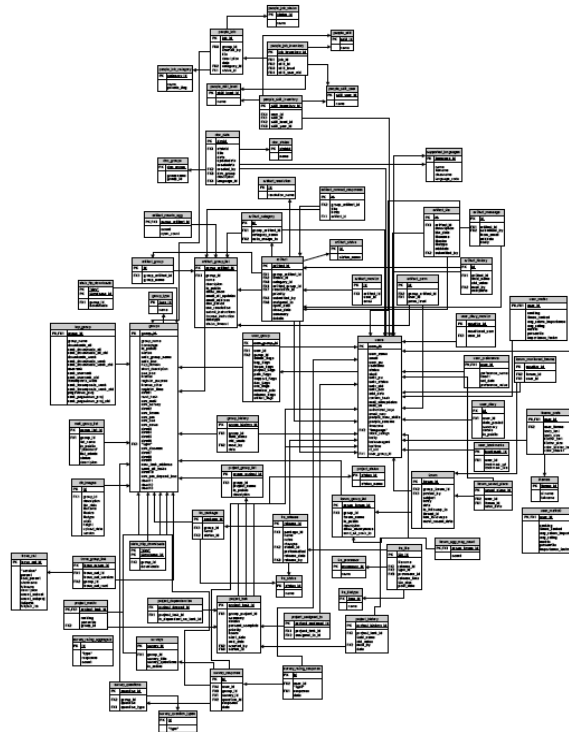
```
Alle Debian-Pakete in »stable«

Zuletzt geändert: Tue, 10 Jul 2007 08:06:51 +0000
Copyright (c) 1997-2005 SPI;
Unter <URL:http://www.debian.org/license> finden Sie die Lizenzbedingungen.

3270-common (3.3.4p6-3.3) Common files for IBM 3270 emulators and pr3287
3dchess (0.8.1-12) 3D chess for X11
3ddesktop (0.2.9-6) "Three-dimensional" desktop switcher
6tunnel (0.11rc2-2) TCP proxy for non-IPv6 applications
915resolution (0.5.2-9) resolution modification tool for Intel graphic chipset
9base (2-3) Plan 9 userland tools
9menu (1.8-1.1) Creates X menus from the shell
9wm (1.2-7) emulation of the Plan 9 window manager 8-1/2
a2ps (1:4.13b.dfsg.1-1) GNU a2ps - 'Anything to PostScript' converter and pretty-printer
a2ps-perl-ja (1.45-4) perl version of Miguel Santana's a2ps (supports KANJI)
aa3d (1.0-5) ASCII art stereogram generator
aap (1.072-1.1) make-like "expert system" for building software
aap-doc (1.072-1.1) make-like "expert system" for building software (documentation)
aatv (0.3-3) A program to watch TV in a text-based console
abakus (0.91-1) calculator for KDE
abcde (2.3.99.6-1) A Better CD Encoder
abcm2ps (4.12.18-1) Translates ABC music description files to PostScript
abcmidi (20060422-1) converter from ABC to MIDI format and back
abcmidi-yaps (20060422-1) yet another ABC to PostScript converter
abichk (1.2-2) binary compatibility checking tool
abiword (2.4.6-1.1) WYSIWYG word processor based on GTK2
abiword-common (2.4.6-1.1) WYSIWYG word processor based on GTK2
abiword-gnome (2.4.6-1.1) WYSIWYG word processor based on GTK2/GNOME2
abiword-help (2.4.6-1.1) online help for AbiWord
abiword-plugins (2.4.6-1.1) plugins for AbiWord
abiword-plugins-gnome (2.4.6-1.1) plugins for AbiWord (with GNOME dependency)
abtex (0.9-beta2-2) LaTeX class for writing documents in ABNT standard
abook (0.5.6-3) text-based ncurses address book application
aboot (0.9b-3) Linux bootloader for the SRM console
aboot-base (0.9b-3) base files required for bootable media on Linux/Alpha
aboot-cross (0.9b-3) utility to create bootable ISO-Images for Linux/Alpha
aboot-installer (0.0.25) [debian-installer] Install aboot on a hard disk
abs-guide (4.1-1) [non-free] The Advanced Bash-Scripting Guide
abuse (1:0.7.0-5) SDL port of the Abuse action game
abuse-frabs (2.10-7) levels and graphics for Abuse
abuse-lib (2.00-18) original levels for Abuse
abuse-sdl (1:0.7.0-5) SDL port of the Abuse action game
abuse-sfx (2.00-9) [non-free] sound effects for Abuse
acct (6.4-pre1-3) The GNU Accounting utilities for process and login accounting
ace-of-penguins (1.2-8) Solitaire-games with penguin-look
acfax (981011-11) Receive faxes using your radio and sound card
acheck (0.5.1) Check common localisation mistakes
acheck-rules (0.3.1) Basic rules for acheck
```


## IV.c) Sourceforge.net list

- hosted by University Notre Dame
- monthly data ~25 GB
- SQL-queries via webinterface



# IV.d) Gathering information – Part I

1<sup>st</sup>: look at the project website



The screenshot shows the Samba project website. At the top, there is a yellow search bar and a navigation menu. Below that is a blue banner with the Samba logo and the tagline "Opening Windows to a Wider World". The main content area is divided into several sections: "think samba", "get samba", "learn samba", "talk samba", and "hack samba". Each section contains links to various resources. A prominent blue banner highlights "10 years of Samba", which contains a paragraph about the project's history and a section titled "The first version". Below that, there is a section for "nbserver 1.5" and "nbserver 1.5.20".

**think samba**

- [What Is Samba?](#)
- [Latest News](#)
- [FAQ](#)

**get samba**

- [Download Info](#)
- [Binaries](#)
- [How To Install](#)
- [GUIs](#)

**learn samba**

- [Official HOWTO](#)
- [By Example](#)
- [Using Samba](#)
- [Docs and Books](#)

**talk samba**

- [List Subscribe](#)
- [List Archives](#)
- [IRC](#)
- [Etiquette](#)

**hack samba**

- [Devel Overview](#)
- [SVN Source](#)
- [Historical CVS Source](#)
- [Build Farm](#)
- [Bug Reports](#)

**contact samba**

- [Samba Team](#)

**10 years of Samba!**

Who would have believed that [this announcement](#) would eventually lead to the project that is now Samba? Thanks to the involvement of countless individuals over the years Samba has come a long way in the last decade, and is set to continue improving for the foreseeable future.

**The first version**

The first tentative version was written in the first few days of 1992. You can read the gory details at <http://ftp.samba.org/ftp/unpacked/samba/docs/history> or you can grab that very first version from <http://ftp.samba.org/ftp/samba/old-versions/server-05.tar.Z>.

That first version still brings back happy memories, and some cringing at the awful code. It was my first sockets program, and it really shows. Last year while Jeremy and I were at a conference we downloaded that first version and managed to coax NT4 into getting a directory listing out of it, but apart from that it really is quite useless with current clients. The protocol has changed too much.

**nbserver 1.5**

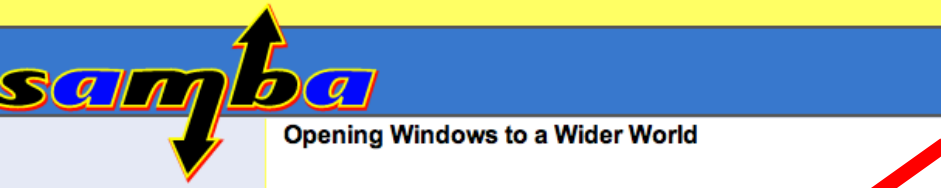
After a long break, nbserver 1.5 was released in December 1993. This version saw the addition of "nmbserver" (now known as nmbd) and smbclient and the renaming of the primary daemon to "smbserver". This is the first version that was influenced by a SMB specification, the earlier versions were based purely on sniffer dumps. This version was also the first to sport the ability to act as a print server.

The nbserver 1.5 was released under the GNU General Public License, unlike earlier version which were released under a "do what you like" style of license. This largely reflects the fact that I had become a huge fan of Linux and was impressed by the progress being made by the Linux community. I've never regretted this decision, and I think the GPL has been very good for Samba.

**nbserver 1.5.20**

# IV.e) Gathering information – Part II

1<sup>st</sup>: look at the project website



- think samba**
  - [What Is Samba?](#)
  - [Latest News](#)
  - [FAQ](#)
- get samba**
  - [Download Info](#)
  - [Binaries](#)
  - [How To Install](#)
  - [GUIs](#)
- learn samba**
  - [Official HOWTO](#)
  - [By Example](#)
  - [Using Samba](#)
  - [Docs and Books](#)
- talk samba**
  - [List Subscribe](#)
  - [List Archives](#)
  - [IRC](#)
  - [Etiquette](#)
- hack samba**
  - [Devel Overview](#)
  - [SVN Source](#)
  - [Historical CVS Source](#)
  - [Build Farm](#)
  - [Bug Reports](#)
- contact samba**
  - [Samba Team](#)

## 10 years of Samba!

Who would have believed that [this announcement](#) would countless individuals over the years Samba has come a l

### The first version

The first tentative version was written in the first few days <http://ftp.samba.org/ftp/unpacked/samba/docs/history> or y <http://ftp.samba.org/ftp/samba/old-versions/server-05.tar.Z>

That first version still brings back happy memories, and s year while Jeremy and I were at a conference we downlo but apart from that it really is quite useless with current cli

### nbserver 1.5

After a long break, nbserver 1.5 was released in Decemb smbclient and the renaming of the primary daemon to "s versions were based purely on sniffer dumps. This versio

The nbserver 1.5 was released under the GNU General F style of license. This largely reflects the fact that I had bec community. I've never regretted this decision, and I think t

### nbserver 1.5.20

Andrew Tridgell [Profil anzeigen](#)

For all those of you who have pathworks for DOS (TCP/IP) but want to be able to use file services from servers other than DecStations and VAXes this may be a solution for you.

I finally got sick of not being able to access my Sun's disk using Pathworks and have now written a server which will compile and run on many unix boxes. I have tested it on a Sun Sparcstation, and have compiled but not properly tested it on a RS6000 and a Sequent symmetry. With this server installed I can now mount file services from my Sun and have them function just like those mounted from a decstation. No software changes are required at the PC end.

If anyone is interested in getting a copy of the code they should send mail to me at "Andrew.Tridg...@anu.edu.au" and I will send them a uuencoded copy of the source. I should warn that this should be regarded as version 0.5, as there are some minor problems that remain to be fixed.

The program is written in C, and will compile with the standard Ultrix, SunOS, AIX or DYNIX compilers. The code is only sparsely commented and the protocol that it implements is an absolute nightmare so if you want to make changes then good luck!

Note that I wrote the server without reference to any source of the dec server, and without disassembling the server and without any inside information so I presume that nobodys rights have been infringed.

Cheers!

Andrew

-----  
 Andrew Tridgell                      CSLab, Research School of Physical Sciences  
 Andrew.Tridg...@anu.edu.au      Australian National University (x2454)  
 -----

## IV.f) Sending e-mail – Part II

2<sup>nd</sup>: write an e-mail:

Hello <Name of Developer>,

I'm studying the reasons why Open Source Software projects get founded. In the last couple of days I've been studying your project. I'm almost done with gathering data, I just want to ask if you can help me with some concluding questions.

<Listing of all manually gathered data>

Is this information correct?

## IV.g) Sending e-mail (cont.)

Can you briefly describe why you started the project?

Approx. when did you have the idea for your project? (E.g. August 2006)

Approx. when did you start your project?

Approx. when did you publish your project?

Approx. when did you publish your first release?

(here only which data is still not known from the website)

How many developers are actually involved in your project?

What kind of tools do you use in your project to communicate respectively share

source code, documentation, etc.?

Information will be made anonymous for this study.

Please let me know if you have question and thanks for helping me out,



For all projects (84):

<b>Reasons for founding</b>	<b># of projects</b>	<b>Percent</b>
scratching an itch	69	82.14%
exercise	7	8.33%
curiosity	6	7.14%
research	5	5.95%
spinoff	3	3.57%
hobby	2	2.38%
fundamental reason	1	1.19%
other	2	2.38%
no data	2	2.38%



- 69 projects founded by 'Scratching an itch'
    - ▶ projects seems to be founded by a need of a person
  - 39 of these have one developer
  - over all projects 47 projects have only one developer
    - ▶ only 8 projects have no 'Scratching an itch' as reason
  - Nearly the half of the projects 'Scratching an itch' have only one developer
- ▶ Majority of projects are founded and maintained by one person to satisfy their needs
- ▶ Raymond is (almost) right with his thesis

### **Correlation Starting Points and Reasons for Founding**

- 'Scratching an itch'
  - from idea to start the project: average 1.4 months
  - start of the project to 1st release: average 3.6 months
- another reason for founding
  - from idea to start the project: average 7.2 months
  - start of the project to 1st release: average 2.8 months
- data not very detailed
  - ▶ outliers have significant impact on the results
  - larger database to absorb outliers

- no direct cover between more collaboration means when more than one developer is involved in a project
- 45% of all projects use a revision control system
- Mailing list are used by about 25% of all projects
- No tools are used by about 19% of all projects

- Sourceforge:
  - 14 of 16 projects have 'Scratching an itch' as reason
  - 10 of 16 projects have one developer
    - whereby 9 projects have 'Scratching an itch' as reason
- Debian:
  - 22 of 26 projects have 'Scratching an itch' as reason
  - 12 of 26 projects have one developer
    - whereby 12 projects have 'Scratching an itch' as reason
- Both:
  - 32 of 42 projects have 'Scratching an itch' as reason
  - 18 of 42 projects have one developer
    - whereby 18 projects have 'Scratching an itch' as reason

- Main reason for founding OSS projects in this study is 'scratching an itch'
- But this is not the only reason
- Other reasons are not as distinct as 'scratching an itch', but there are still common reasons
  
- Only scratched on the surface
- Deeper, more detailed and wider spread survey might confirm the results and give deeper insight into the topic

# Thank you!

# Any further questions?