



# State of SciPy

Jarrold Millman  
Travis N. Vaught

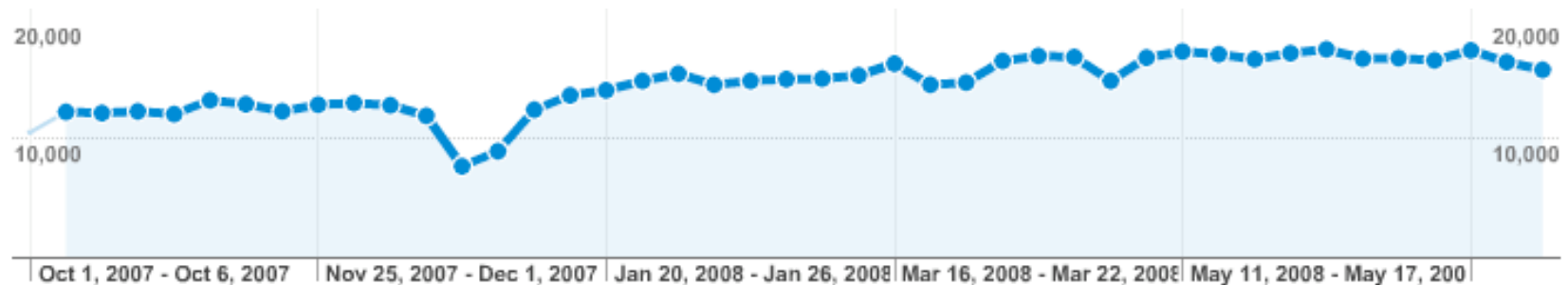
August 21, 2008

# Overview

- Introduction (community statistics)
- Past: Where have we been
- Present: Where are we
- Future: Where are we going?

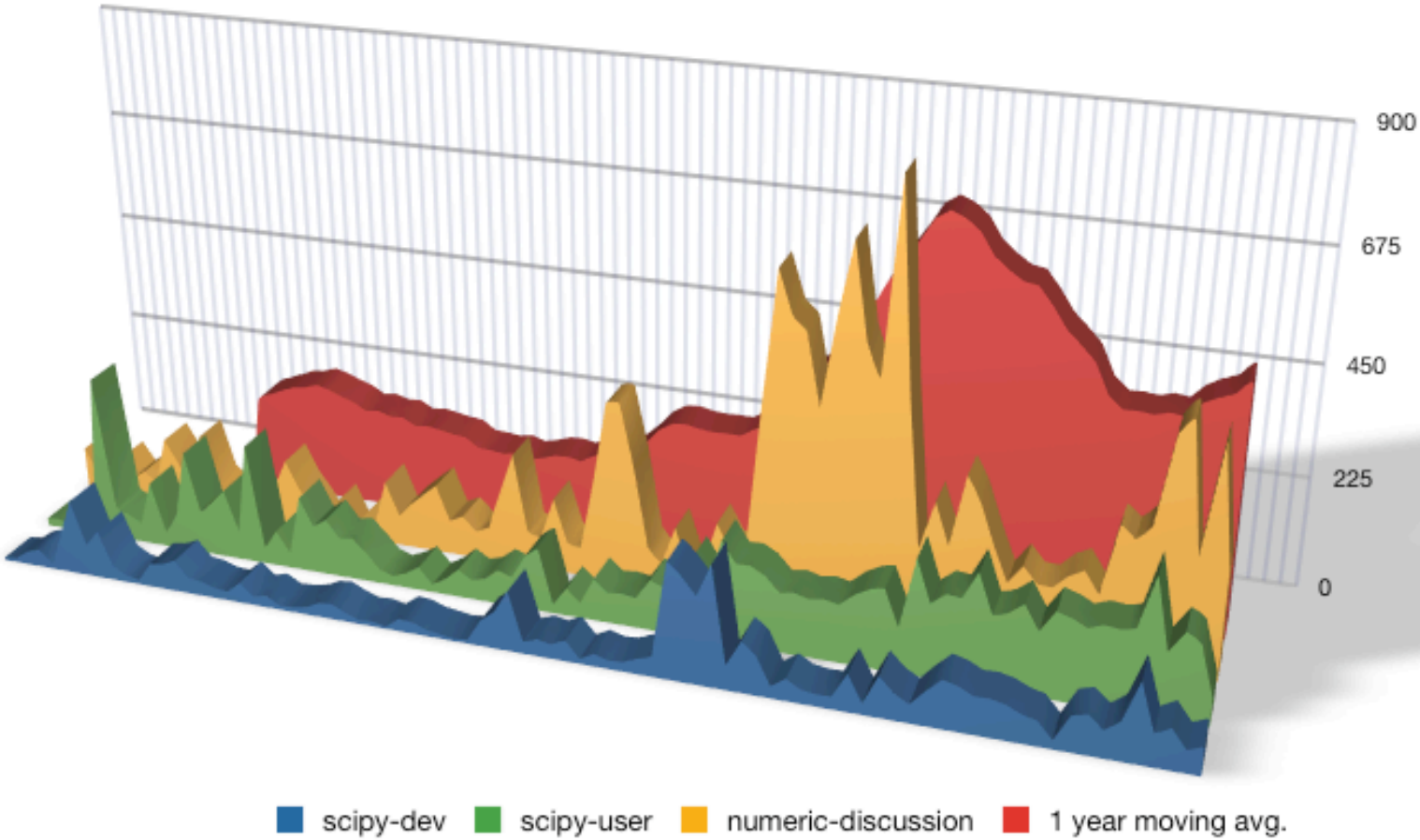
# scipy.org Activity

- There has been marked growth in the community of users around scientific computing in Python in the past year.



# Mailing List Activity

Monthly Mailing List Traffic (KBytes) August 2001 - August 2008

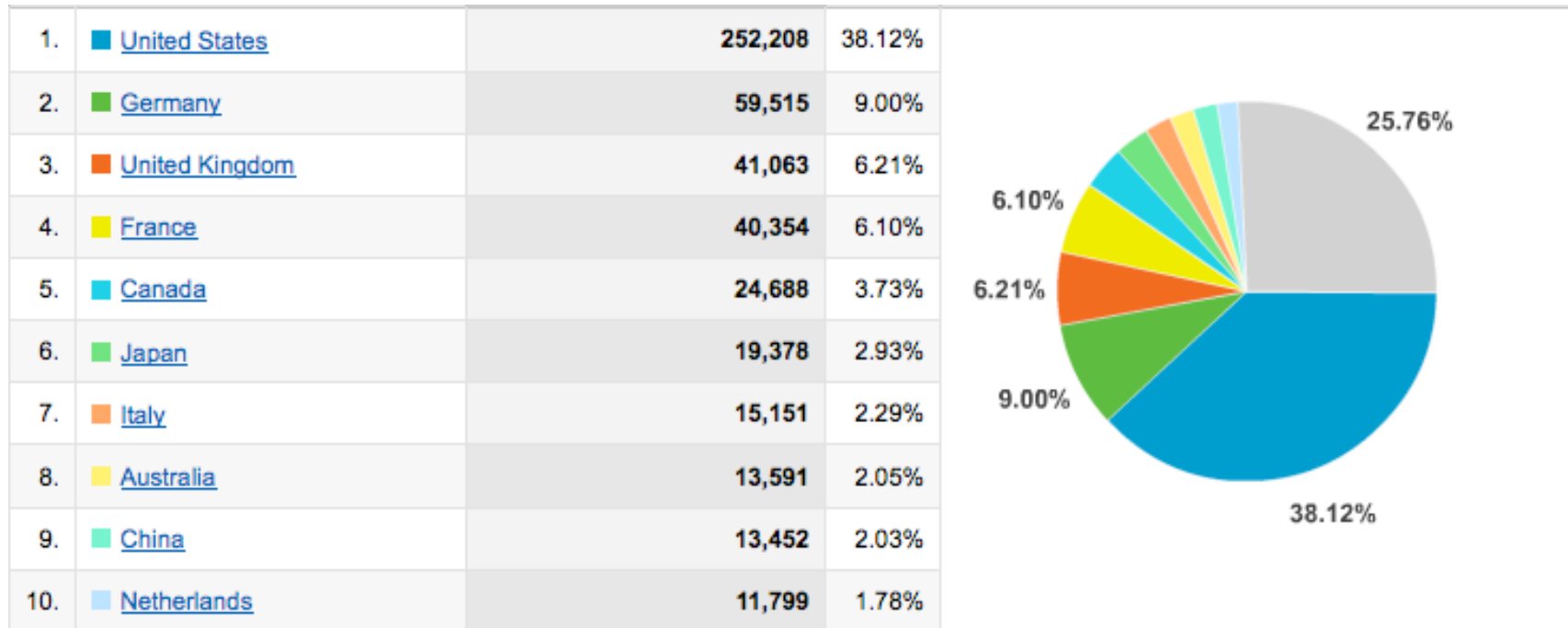


# The SciPy Community is Global



# scipy.org Usage by Country

- Visits to scipy.org since Sept. 2008



# Student Sponsorship

David Cournapeau	Japan
Stefan van der Walt	South Africa
Rahul Garg	Alberta
Seshadri Tirunillai	USC
Amir Khosrowshahi	Berkeley
Dav Clark	Berkeley
Gabriel Gellner	Toronto
Min Ragan-Kelley	Berkeley
Ryan May	OU
David Warde-Farley	Toronto

Many thanks to the sponsoring organizations!



# Short History

- 1994 — Python Matrix object (Jim Fulton)
- 1995 — Numeric born (Jim Hugunin, Konrad Hinsen, Paul Dubois, David Ascher, Jim Fulton)
- 2000 — Numeric moves to sourceforge (Project registered as numpy)
- 2001 — SciPy born (Pearu Peterson, Travis Oliphant, Eric Jones)
- 2001 — IPython born (Fernando Perez)
- 2002 — SciPy '02 - Python for Scientific Computing Workshop
- 2003 — matplotlib born (John Hunter)
- 2003 — Numarray (Perry Greenfield, J. Todd Miller, Rick White, Paul Barrett)
- 2006 — NumPy 1.0 Released



# Code Swarm

- Now for something completely different

# NumPy Book Announcement

Trelgol

http://www.tramy.us/

Getting Started Latest Headlines

## Trelgol Publishing

Setting information free



0110101

### Products

Name	Author	Unit Price	Buy Now	Total Time	May Sell Rights	World Price
<a href="#">Guide to NumPy</a> you will receive the current version and updates. More purchasing options are available <a href="#">here</a> .	Travis E. Oliphant	\$39.99 <sup>1</sup>	 	3 yrs Aug 21, 2008	N	\$200,000 <sup>2</sup>
Generic layout file for writing books with LyX (this is just a formatting file useful if you want to write another book using LyX)	Travis E. Oliphant	\$5.00		3 yrs Aug 21, 2008	N	\$500

<sup>1</sup> You may still donate this amount, but the book is now in the public domain.  
<sup>2</sup> The final actual sold "world price" was about \$105,000. Much of this money went to fund students.

### World Price Briefly Explained

This represents the total price the author requires before freely releasing the content. No one is required to individually pay this amount.

Done

410 jarrod.millman@gmail.com

# Present: Where we are

- NumPy/SciPy packages
- Community involvement
- Larger ecosystem

# 2007-2008 NumPy/SciPy Releases

- SciPy 0.6.0 (September 2007)
- NumPy 1.0.4 (November 2007)
- NumPy 1.1.0 (May 2008)
- NumPy 1.1.1 (August 2008)

# Major Changes

- New MaskedArrays
- IO code overhaul
- Improved test coverage
- Numscscons
- Focus on coding style

# Release Management

- Better packaging
- Release numbering
- Buildbot

# Community Involvement

- Blogs (<http://planet.scipy.org>)
- IRC (`#scipy` on freenode)
- Mailing lists
- Sprints
- Conferences

# List Traffic

- Average daily message count:
  - numeric-discussion, scipy-user, scipy-dev: 44 messages/day
  - Adding in ipython and matplotlib yields another 15 per day
  - If you monitor enthought-dev you'll be reading another 31 per day
- Total: 90 messages per day
- Yes, it's hard to keep up.



# Sprints

- August 2007 — SciPy 2007 sprint
- December 2007 — SciPy sprint at UCB
- February 2008 — Sage days 8 at Enthought
- March 2008 — NIPY/IPython sprint in Paris
- April 2008 — SciPy sprint at UCB
- July 2008 — MayaVi sprint at Enthought

# Conferences

- PyCon 2008
- 2008 SIAM meeting
- EuroSciPy 2008
- SciPy 2008

# The Larger Ecosystem

- Major releases
- Distribution
- New Tools

# Major Releases

- matplotlib 0.98
- ETS 3
- MayaVi 2

# Distribution

- Python Package Index
- Python(x,y)
- EPD
- Sage

# Future: Where are we going?

- NumPy/SciPy packages
- The Larger Ecosystem
- Getting involved

# NumPy/SciPy releases

- Upcoming releases
- Release management

# NumPy 1.2

- Sphinx-based documentation!!
- Nose-based testing framework (GSoC project)
- Changes to histogram and median
- Arrayiterators



# NumPy 1.2 open issues

- ABI compatibility
- Generalized UFuncs
- C-API cleanups
- Release date

# SciPy 0.7

- Sparse matrices improvements
- New functionality: Radial basis functions, hierarchical clustering, constants
- Nose-based testing framework
- Sandbox removed

# Release Management

- Time-based releases
- Code Review
- Enhancement Proposals
- Distributed Version Control

# Time- vs. Feature-based Releases

## **Feature-based**

A release cycle under this model is driven by deciding what features will be in the next release. Once all the features are complete, the code is stabilized and finally a release is made. Obviously this makes it relatively easy to predict what features will be in the next release, but extremely difficult to determine when the release will occur.

## **Time-based**

A release cycle under this model is driven by deciding when the next release will be. This, of course, makes predicting when the release will be out extremely easy, but makes it difficult to know exactly what features will be included in the release.

# Time-based Release Issues

- Branching
- Reviewing
- Testing
- Reverting
- Postponing
- Releasing

# Time-based Releases: Branching

In order to be able to release on schedule requires that the mainline of development (the trunk) is extremely stable. This requires that a significant amount of work be conducted on branches.

# Time-based Releases: Reviewing

Another important way to improve the quality of project and keep the trunk in shape is to require peer code review and consensus among the core developers on which branches are ready to be merged.

# Time-based Releases: Testing

A full test suite is also essential for being able to regularly release code.



# Time-based Releases: Reverting

Sticking to release schedule requires occasionally reverting commits.

# Time-based Releases: Postponing

It also requires postponing branch merges until the branch is ready for release.

# Time-based Releases: Releasing

Since time-based release management relies on a regular releases, the cost of making a release needs to be minimized. In particular, we need to make it much, much easier to create the packages, post the binaries, create the release notes, and send out the announcements.

# Release Management Practices

- Code review
- Enhancement proposals
- Distributed version control

# Code Review: Rietveld Tool

Rietveld Code Review Tool

[Help](#) | [Bug tracker](#) | [Discussion group](#) | [Source code](#) | [Sign in](#)<sup>(2)</sup>

**Issues**   Repositories

[Recent Issues](#) | [Sign in](#) with your [Google Account](#) to create issues and add comments

## Issue [2946](#): Update NumPy API format to support updates that don't break binary compatibility

Can't Edit  
Can't Publish+Mail  
[Start Review](#)

**Created:**  
5 days, 1 hour ago by [Stefan](#)

**Modified:**  
1 day, 6 hours ago

**Reviewers:**  
[Andrew Straw](#)

**SVN Base:**  
<http://svn.scipy.org/svn/numpy/trunk/>

### ▼ Description

The current NumPy API number, stored as NPY\_VERSION in the header files, needs to be incremented every time the NumPy C-API changes. The counter tells developers with exactly which revision of the API they are dealing. NumPy does some checking to make sure that it does not run against an old version of the API. Currently, we have no way of distinguishing between changes that break binary compatibility and those that don't.

The proposed fix breaks the version number up into two counters -- one that gets increased when binary compatibility is broken, and another when the API is changed without breaking compatibility.

Backward compatibility with packages such as Matplotlib is maintained by renaming NPY\_VERSION to NPY\_BINARY\_VERSION.

### ▼ Patch Set 1

Total comments: 4

Created: 5 days, 1 hour ago [Download raw patch set](#)

Raw unified diffs	Stats	Side-by-side diffs with inline comments	Delta from patch set	Patch
<a href="#">numpy/core/code_generators/generate_numpy_api.py</a>	1 chunk 20 lines	0 comments	<a href="#">2</a>	<a href="#">Download</a>
<a href="#">numpy/core/include/numpy/ndarrayobject.h</a>	1 chunk 31 lines	4 comments	<a href="#">2</a>	<a href="#">Download</a>
<a href="#">numpy/core/src/multiarraymodule.c</a>	1 chunk 13 lines	0 comments	<a href="#">2</a>	<a href="#">Download</a>

### ▼ Patch Set 2 : Updated with Andrew's suggestions.

Created: 4 days, 16 hours ago [Download raw patch set](#)

Raw unified diffs	Stats	Side-by-side diffs with inline comments	Delta from patch set	Patch
-------------------	-------	---	----------------------	-------

# Code Review: Rietveld Tool, cont.

```
25
26 /* There are several places in the code where an array of dimensions is
27 * allocated statically. This is the size of that static allocation.
28 *
29 * The array creation itself could have arbitrary dimensions but
30 * all the places where static allocation is used would need to
31 * be changed to dynamic (including inside of several structures)
32 */
33
34 #define NPY_MAXDIMS 32
35 #define NPY_MAXARGS 32
36
37 /* Used for Converter Functions "O&" code in ParseTuple */
38 #define NPY_FAIL 0
39 #define NPY_SUCCEED 1
40
41 /* Helpful to distinguish what is installed */
42 #define NPY_VERSION 0x0100000B

```

```
43
44 /* Some platforms don't define bool, long long, or long double.
45 Handle that here.
46 */

```

```
25
26 /* There are several places in the code where an array of dimensions is
27 * allocated statically. This is the size of that static allocation.
28 *
29 * The array creation itself could have arbitrary dimensions but
30 * all the places where static allocation is used would need to
31 * be changed to dynamic (including inside of several structures)
32 */
33
34 #define NPY_MAXDIMS 32
35 #define NPY_MAXARGS 32
36
37 /* Used for Converter Functions "O&" code in ParseTuple */
38 #define NPY_FAIL 0
39 #define NPY_SUCCEED 1
40
41 /* Helpful to distinguish what is installed */
42 /* NPY_VERSION is no longer used -- see NPY_BINARY_VERSION below. */
43 #define NPY_VERSION 0x1000000B

```

Andrew Straw 2008/08/16 07:20:56 there's no need to change the actual value of NPY\_

Stefan 2008/08/16 07:33:19 On 2008/08/16 07:20:56, Andrew Straw wrote: > ther

44

Andrew Straw 2008/08/16 07:20:56 Perhaps move NPY\_VERSION below the NPY\_BINARY\_VERS

Stefan 2008/08/16 07:33:19 On 2008/08/16 07:20:56, Andrew Straw wrote: > Perh

```
45 /* Whenever (NPY_BINARY_VERSION & NPY_BINARY_COMPATIBLE) changes,
46 external libraries that depend on NumPy need to be recompiled.
47
48 In other words, NPY_BINARY_VERSION consists of two parts:
49
50 0xMMMMKKKK
51
52 where M gets increased whenever binary compatibility is broken,
53 and K gets increased whenever the API is changed but binary
54 compatibility is maintained.
55 */
56 #define NPY_BINARY_VERSION 0x00010001
57 #define NPY_BINARY_COMPATIBLE 0xFFFF0000

```

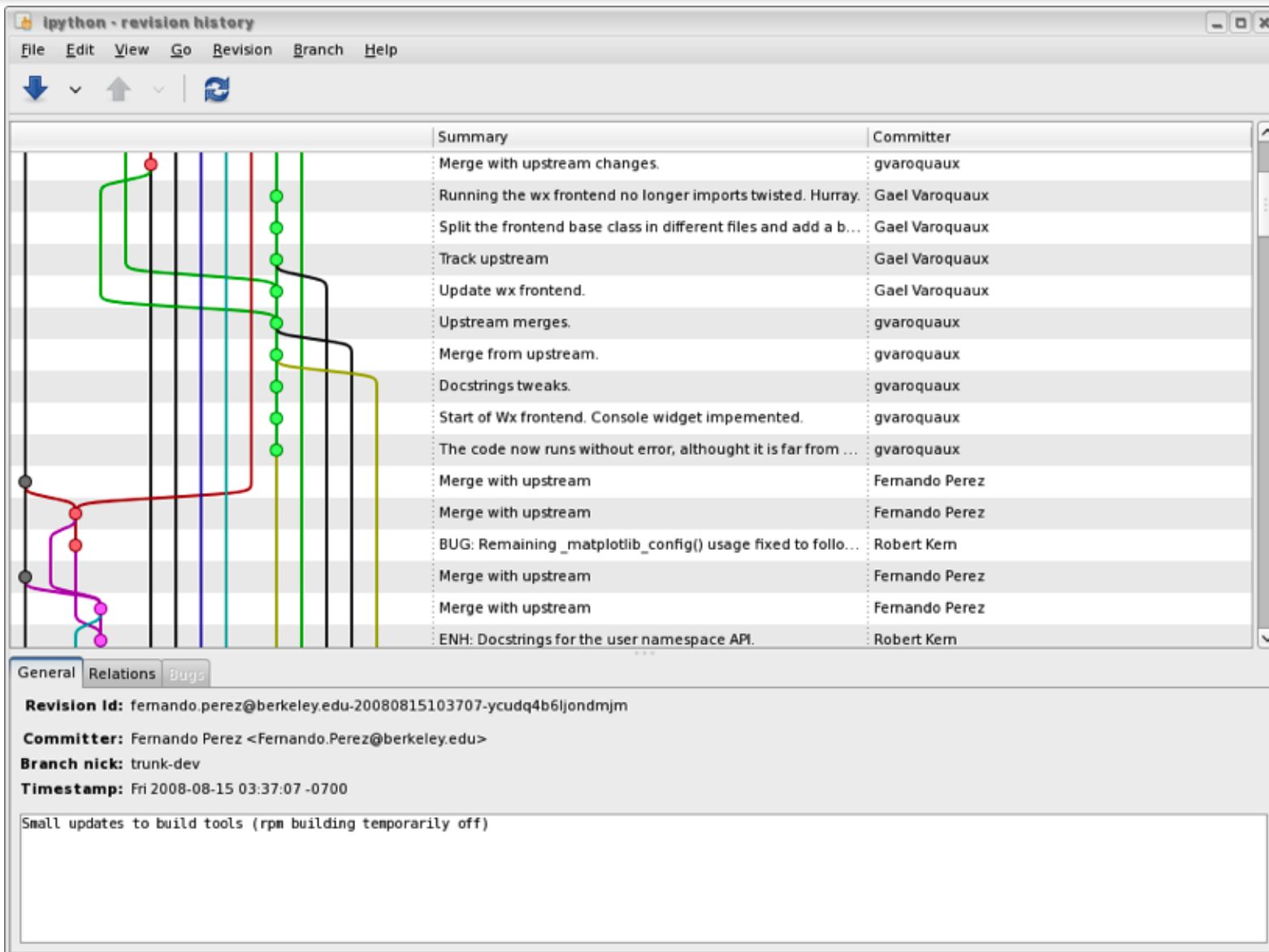
```
58
59 /* Some platforms don't define bool, long long, or long double.
60 Handle that here.
61 */

```

# Enhancement Proposals

- A Simple File Format for NumPy Arrays
- Implementing date/time types in NumPy
- Matrix Indexing
- Runtime Optimization
- Solvers Proposal

# Distributed Version Control



The screenshot shows the 'ipython - revision history' window. On the left, a commit graph displays various branches (black, red, green, blue, cyan, yellow) and their relationships. On the right, a table lists revisions with their summaries and committers.

Summary	Committer
Merge with upstream changes.	gvaroquaux
Running the wx frontend no longer imports twisted. Hurray.	Gael Varoquaux
Split the frontend base class in different files and add a b...	Gael Varoquaux
Track upstream	Gael Varoquaux
Update wx frontend.	Gael Varoquaux
Upstream merges.	gvaroquaux
Merge from upstream.	gvaroquaux
Docstrings tweaks.	gvaroquaux
Start of Wx frontend. Console widget impemented.	gvaroquaux
The code now runs without error, although it is far from ...	gvaroquaux
Merge with upstream	Fernando Perez
Merge with upstream	Fernando Perez
BUG: Remaining _matplotlib_config() usage fixed to follo...	Robert Kern
Merge with upstream	Fernando Perez
Merge with upstream	Fernando Perez
ENH: Docstrings for the user namespace API.	Robert Kern

**General** | Relations | Bugs

**Revision Id:** fernando.perez@berkeley.edu-20080815103707-ycudq4b6ljondmjm

**Committer:** Fernando Perez <Fernando.Perez@berkeley.edu>

**Branch nick:** trunk-dev

**Timestamp:** Fri 2008-08-15 03:37:07 -0700

Small updates to build tools (rpm building temporarily off)



# Launchpad

The screenshot shows a web browser window titled "Launchpad" with the URL "https://launchpad.net/". The browser's address bar includes navigation icons and a search engine dropdown set to "Google". The website header features a navigation menu with tabs for "Launchpad Home", "Code", "Bugs", "Blueprints", "Translations", and "Answers". A user profile for "Jarrod Millman" with a "Log Out" link is visible in the top right. The main content area is divided into three columns: "What is Launchpad?", "What's new?", and "Featured projects".

**What is Launchpad?**  
Launchpad is a hosting service for open source projects that's big on collaboration. Take a tour... (new!)

**Start here:**

- Projects
- Distributions
- People and Teams

**What's new?**

- ★ **Bugzilla and Trac plugins now in beta**  
Do you run a Bugzilla or Trac instance? Help beta test our new plugins and share bug comments with Launchpad.
- ★ **Karma for commits**  
You'll now get karma for each commit you make to a branch that's registered in Launchpad.
- ★ **New Launchpad user guide**  
If you're new to Launchpad, our new user guide will take you through each application step by step.
- ★ **Help test Launchpad's new API!**  
Our new internet-services API is now in beta test.

Updated 2008-08-20. More...

**Featured projects**

★ Anewt	🛡️ LottaNZB
📧 Awn	📺 Me TV
📁 Bauble	🔍 MySQL
📦 Bazaar	🖥️ PyRoom
🐧 GNOME Do	🎓 SchoolTool
🗨️ Drizzle	🌊 Silva CMS
🏎️ Extreme Tux Racer	🌐 TangoCMS
🌱 Exaile	🔪 Terminator
🌿 GASP	🐧 Ubuntu
👤 Get-You	🌐 Zope.org

Help improve Launchpad © 2004-2008 Canonical Ltd. | Terms of use

Done launchpad.net 410 jarrod.millman@gmail.com

# Launchpad Code Review

The screenshot shows a web browser window titled "Active code reviews for IPython". The address bar contains the URL "https://code.launchpad.net/ipython/+activereviews". The page features a navigation bar with tabs for "Overview", "Code", "Bugs", "Blueprints", "Translations", and "Answers". The "Code" tab is selected. Below the navigation bar, the page title "Active code reviews for IPython" is displayed in orange. A summary of activity is provided: "25 branches owned by 7 people and 2 teams, 159 commits by 18 people in the last month" and "3 active reviews, 0 approved merges". A section titled "You can browse the source code for the development focus branch or get a copy of the branch using the command:" includes the command `bzr branch lp:ipython`. Below this, a note states "There are download files available for IPython." A table lists three active reviews, each requested by Fernando Perez. The table has four columns: "Branch Merge Proposal", "Requested By", "Date Review Requested", and "Vote Summary".

Getting Started Latest Headlines

launchpad IPython

Overview Code Bugs Blueprints Translations Answers

## Active code reviews for IPython

25 branches owned by 7 people and 2 teams, 159 commits by 18 people in the last month

3 active reviews, 0 approved merges

You can [browse the source code](#) for the development focus branch or get a copy of the branch using the command:

```
bzr branch lp:ipython
```

There are [download files](#) available for IPython.

Branch Merge Proposal	Requested By	Date Review Requested	Vote Summary
lp~laurent-dufrechou/ipython/trunk-dev ⇒ lp:ipython	Fernando Perez		no votes (no comments)
lp~fdo.perez/ipython/trunk-dev ⇒ lp:ipython	Fernando Perez		no votes (no comments)
lp~ipython-dev/ipython/ipython-daemon ⇒ lp:ipython	Fernando Perez		no votes (no comments)

[Help improve Launchpad](#) © 2004-2008 Canonical Ltd. | [Terms of use](#)

Done code.launchpad.net 410 jarrod.millman@gmail.com

# Launchpad

The screenshot shows a web browser window with the address bar displaying `https://code.launchpad.net/~vcs-imports/numpy/trunk`. The page title is `"~vcs-imports/numpy/trunk" branch in Launchpad`. The main content area features a navigation bar with tabs for `Overview`, `Code`, `Bugs`, `Blueprints`, `Translations`, and `Answers`. Below the navigation bar, the branch name `~vcs-imports/numpy/trunk` is displayed in large orange text. To the right of the branch name, it says `Last modified 4 hours ago`. Below the branch name, there are tabs for `Details`, `Merging`, and `Source Code`. The `Details` tab is selected. The main content area is divided into two columns. The left column contains the following information:   
- **Project:** NumPy   
- **Status:** New   
- **Get this branch:** `bzr branch lp:numpy`   
- **Branch format:** Branch format 6   
- **Repository format:** Packs containing knits without subtree support   
- **Import details**   
- **Import Status:** Reviewed   
- This branch is an import of the Subversion branch from `http://svn.scipy.org/svn/numpy/trunk`.   
- The next import is scheduled to run in 44 minutes.   
- Last successful import was 5 hours ago.   
- A list of import logs, each starting with a green checkmark and followed by details such as `Import started 5 hours ago on galapagos and finished 5 hours ago taking two minutes — see the log`.   
- At the bottom of the list, it says `« No branches proposed for merging into this one. »` and `« Not proposed for merging »`.   
The right column contains a box with the following information:   
- `This branch is public`   
- **Owner**   
- `✓ VCS imports`   
- **Subscribers**   
- `No subscribers.`   
At the bottom of the page, there is a status bar with `Done` on the left and `code.launchpad.net` and `410 jarrod.millman@gmail.com` on the right.

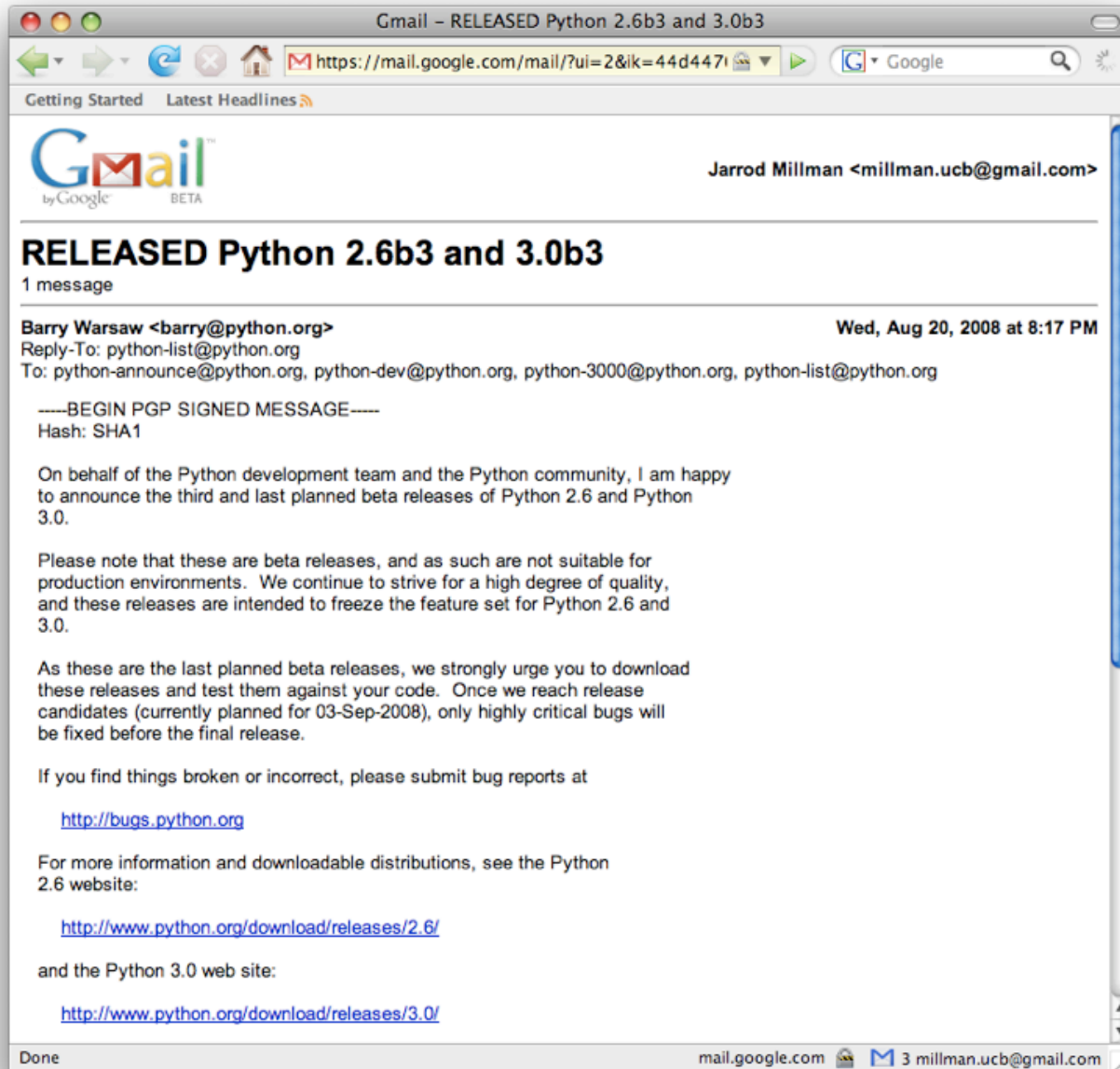
# The Larger Ecosystem

- IPython
- Python

# IPython 0.9

- IPython1 in trunk
- Graphical frontends
- More...

# Python 2.6 and 3.0



# NumPy/SciPy and Python 2.6/3.0

- Improve test coverage
- Port to Python 2.6
- Improve test coverage
- Run 2to3 tool
- Fix issues
- **ISSUE: No 2to3 tool for the C-API!**

# Getting Involved

- Documentation
- Bug-fixes
- Testing
- Code (review, cleanup, contributions)
- Mailing list participation
- Start a local SciPy group
- Code sprints
- Documentation/Bug-fix Days
- Web design



Thank You