PyPy status talk

Samuele Pedroni
Armin Rigo
Antonio Cuni

EuroPython 2009

June 30 2009

Part 1

Becoming complete

PyPy release 1.1

- PyPy released version 1.1 in April 2009

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 2 / 28>

What is PyPy

- An alternate Python virtual machine for C/Posix, for .NET and for Java
- More generally, a way to implement interpreters for various languages

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 3 / 28>

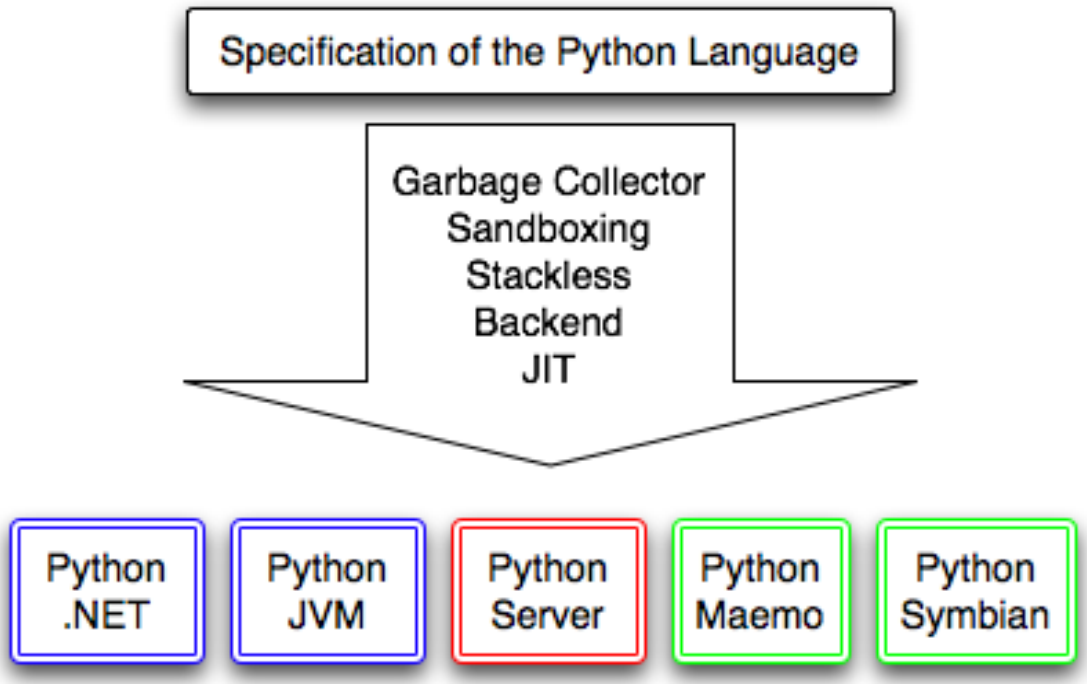
PyPy - motivation

- CPython is nice, but not flexible enough
- IronPython, Jython - bound to their specific VM
- Psyco and Stackless Python are hard to maintain

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 4 / 28>

PyPy: generating a Python Interpreter

- **high level Python specification!**
- layer GCs, JIT, Stackless atop the spec
- **generate interpreters** for targets



Brief history of PyPy

- First sprint in 2003, about 30 more by now
- CPython/Psyco/Jython/Stackless developers participating
- MIT License
- EU Research project 2004-2007
- 2007-now - open source project
- Sponsoring from Google, Nokia
- Soon JIT work cofunded by Germany and Sweden

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 6 / 28>

PyPy 1.1

- Compatible with Python 2.5.2
- Well tested on Windows and Linux 32-bit, but should also work on Mac OS/X and Linux 64-bit
- Runs major packages unmodified, out of the box
- `easy_install` / `distutils` working

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 7 / 28>

PyPy 1.1 - more supported packages

- `ctypes`
- `sqlite`
- `array`, `cPickle`, `cStringIO`, `cmath`, `dbm`, `datetime`, `binascii`...
- `__builtin__`, `__pypy__`, `_codecs`, `_lsprof`, `_minimal_curses`, `_random`, `_rawffi`, `_socket`, `_sre`, `_weakref`, `bz2`, `cStringIO`, `crypt`, `dyngram`, `errno`, `exceptions`, `fcntl`, `gc`, `itertools`, `marshal`, `math`, `md5`, `mmap`, `operator`, `posix`, `pyexpat`, `recparser`, `select`, `sha`, `signal`, `struct`, `symbol`, `sys`, `termios`, `thread`, `time`, `unicodedata`, `zipimport`, `zlib`

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 8 / 28>

PyPy 1.1 - more supported packages

- `ctypes`
- `sqlite`
- `array`, `cPickle`, `cStringIO`, `cmath`, `dbm`, `datetime`, `binascii`...
- `__builtin__`, `__pypy__`, `_codecs`, `_lsprof`, `_minimal_curses`, `_random`, `_rawffi`, `_socket`, `_sre`, `_weakref`, `bz2`, `cStringIO`, `crypt`, `dyngram`, `errno`, `exceptions`, `fcntl`, `gc`, `itertools`, `marshal`, `math`, `md5`, `mmap`, `operator`, `posix`, `pyexpat`, `recparser`, `select`, `sha`, `signal`, `struct`, `symbol`, `sys`, `termios`, `thread`, `time`, `unicodedata`, `zipimport`, `zlib`

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 8 / 28>

PyPy 1.1 - more supported packages

- ctypes
- sqlite
- array, cPickle, cStringIO, cmath, dbm, datetime, binascii...
- __builtin__, __pypy__, _codecs, _lsprof, _minimal_curses, _random, _rawffi, _socket, _sre, _weakref, bz2, cStringIO, crypt, dyngram, errno, exceptions, fcntl, gc, itertools, marshal, math, md5, mmap, operator, posix, pyexpat, recparser, select, sha, signal, struct, symbol, sys, termios, thread, time, unicodedata, zipimport, zlib

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 8 / 28>

PyPy 1.1 - more supported packages

- ctypes
- sqlite
- array, cPickle, cStringIO, cmath, dbm, datetime, binascii...
- __builtin__, __pypy__, _codecs, _lsprof, _minimal_curses, _random, _rawffi, _socket, _sre, _weakref, bz2, cStringIO, crypt, dyngram, errno, exceptions, fcntl, gc, itertools, marshal, math, md5, mmap, operator, posix, pyexpat, recparser, select, sha, signal, struct, symbol, sys, termios, thread, time, unicodedata, zipimport, zlib

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 8 / 28>

PyPy 1.1 - speed

- Performance improved between 10% and 50% since PyPy 1.0
- Python interpreter is now between 0.8 and 2x (and in some corner case 3-4x) slower than CPython
- It starts faster than CPython :-)
- Better garbage collectors (generational and hybrid)
- This is all ignoring the ongoing work on the JIT (part 2 of the talk, not in the 1.1 release)

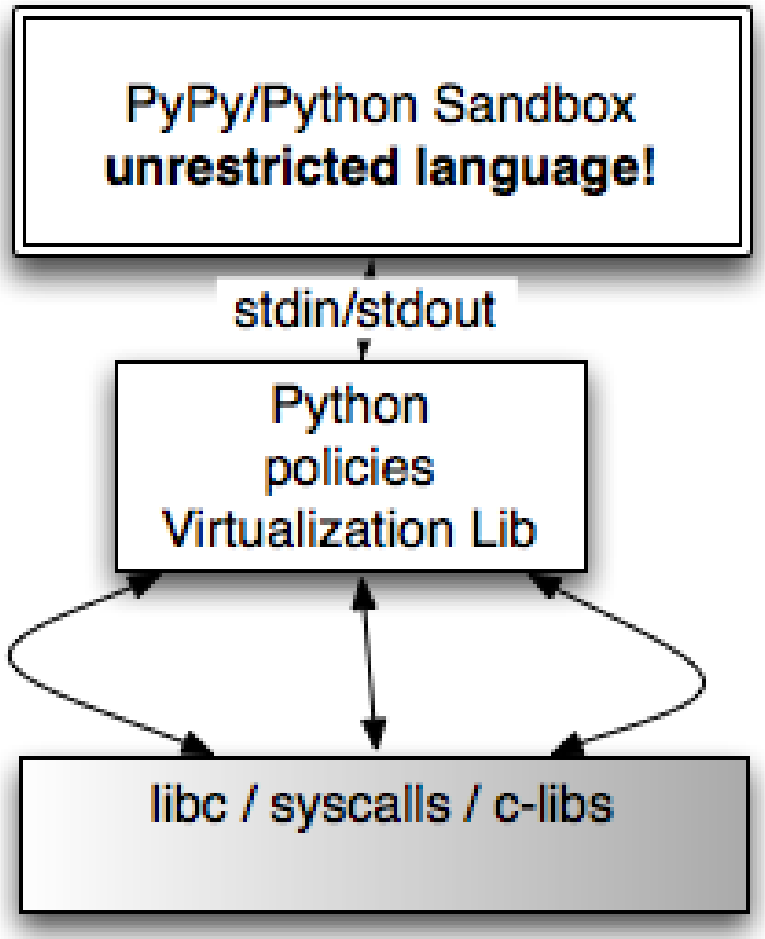
<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 9 / 28>

PyPy 1.1 - sandboxing support

- Sandboxing: “virtualized” version of PyPy, safe to run any untrusted program

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 10 / 28>

PyPy 1.1 - sandboxing support



PyPy 1.1 - sandboxing support

- It is a special version of PyPy that does not do any I/O
- Communicates purely on stdin/stdout
- Run by an outer controlling process (e.g. a CPython or a regular PyPy)
- Really safe by construction!

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 12 / 28>

PyPy 1.1 - more news

- Stackless support complete (tasklet, frame pickling, greenlet)
- Classic classes on by default
- More memory-efficient (e.g. class instances are often only 50% of the size of CPython)
- `_lsprof` profiling

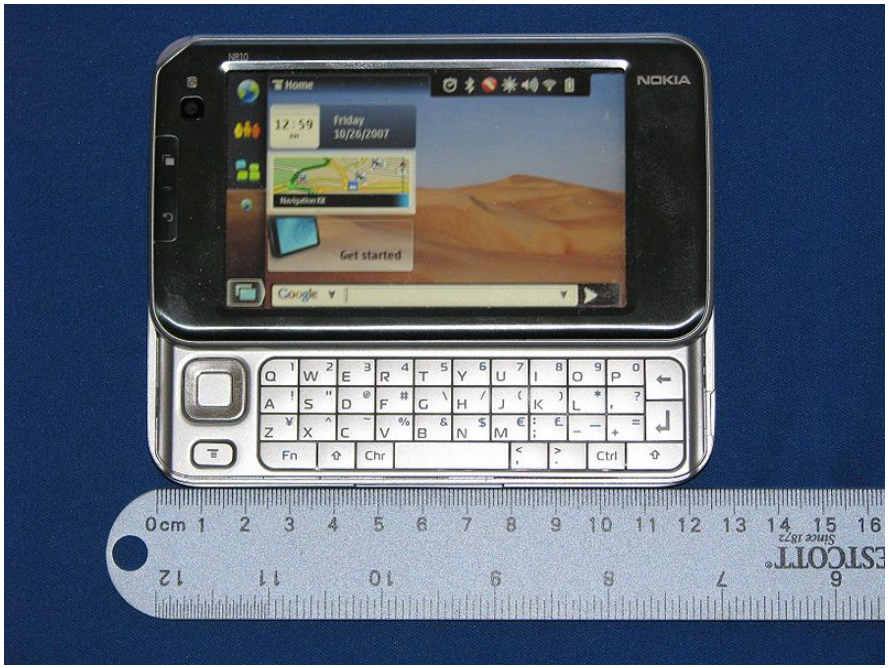
PyPy 1.1 - on CLR/.NET

- PyPy runs on the CLR and can now interface with .NET libraries (more work needed, though)

- PyPy runs on JVM too, but no integration there
- Contributors wanted!

PyPy 1.1 - on Maemo

- Cross-compliation: N810 Internet Tablet by Nokia



- Security, RAM usage, sharing interpreter state across processes... PyPy's approach is good (not fully implemented yet)

What we can run

- We worked a lot on running existing applications on top of PyPy
- Sometimes requiring to change applications slightly
- Especially refcounting details tend to be a problem

```
open('xxx', 'w').write('stuff')
```

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 16 / 28>

CTypes

- Part of CPython stdlib since 2.5
- Official way to have bindings to external (C) libraries for PyPy
- Can handle i.e. `pysqlite-ctypes`, `pyglet`, `pymunk` or `Sole Scion`, almost whatever...
- Contribution to original `ctypes` (better errno handling, bugfixes, tests...)
- Part of Google sponsoring
- Note: a bit slow

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 17 / 28>

Sqlite

- Part of CPython stdlib since 2.5
- We use Gerhard Haering's CTypes version
- Works reasonably well after some fixes
- Included in PyPy

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 18 / 28>

Django

- We run unmodified Django 1.0
- Only sqlite DB backend for now

http://www.djangoproject.com
http://code.djangoproject.com/wiki/DjangoAndPyPy

Pylons

- Worked almost out of the box once eggs were working (1 day)
- No SQLAlchemy yet, obscure problems ahead
- Unmodified, it passes all tests

http://pylonshq.com/

Twisted & Nevow

- Twisted works (60/4500 tests failing)
- Nevow works
- We don't support PyCrypto nor PyOpenSSL and we won't anytime soon (unless someone contributes a CTypes or RPython version)

http://twistedmatrix.com/

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 21 / 28>

Other software

- Anything written in pure Python should just work :-)
- BitTorrent, PyPy translation toolchain, py lib, SymPy, Pinax...
- Various smaller things, templating engines...

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 22 / 28>

Obscure details that people rely on

- Non-string keys in __dict__ of types
- Exact naming of a list comprehension variable
- Relying on untested and undocumented private stuff
- Exact message matching in exception catching code
- Refcounting details

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 23 / 28>

Lessons Learned

- Lessons Learned: There is No Feature Obscure Enough for people not to rely on
- The `pypy-c` interpreter is probably far more compatible to CPython 2.5 than Jython or IronPython
- Main blocker for running apps is missing external modules

<antocuni, pedronis, arigo (EuroPython 2009) PyPy status talk June 30 2009 24 / 28>

Contact / Q&A

PyPy: http://codespeak.net/pypy
Blog: http://morepypy.blogspot.com