Rosetta Benchmark Server

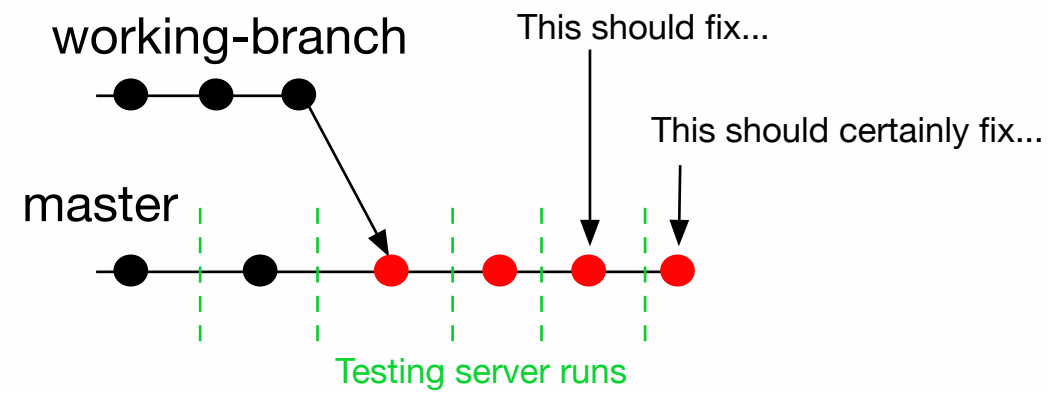## features overview and future directions

Sergey Lyskov   GrayLab@JHU

Pro-active testing workflow
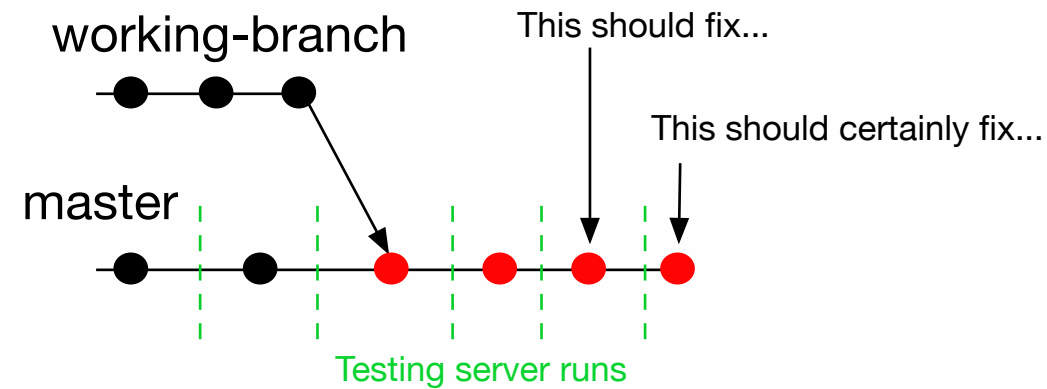
Pro-active testing workflow
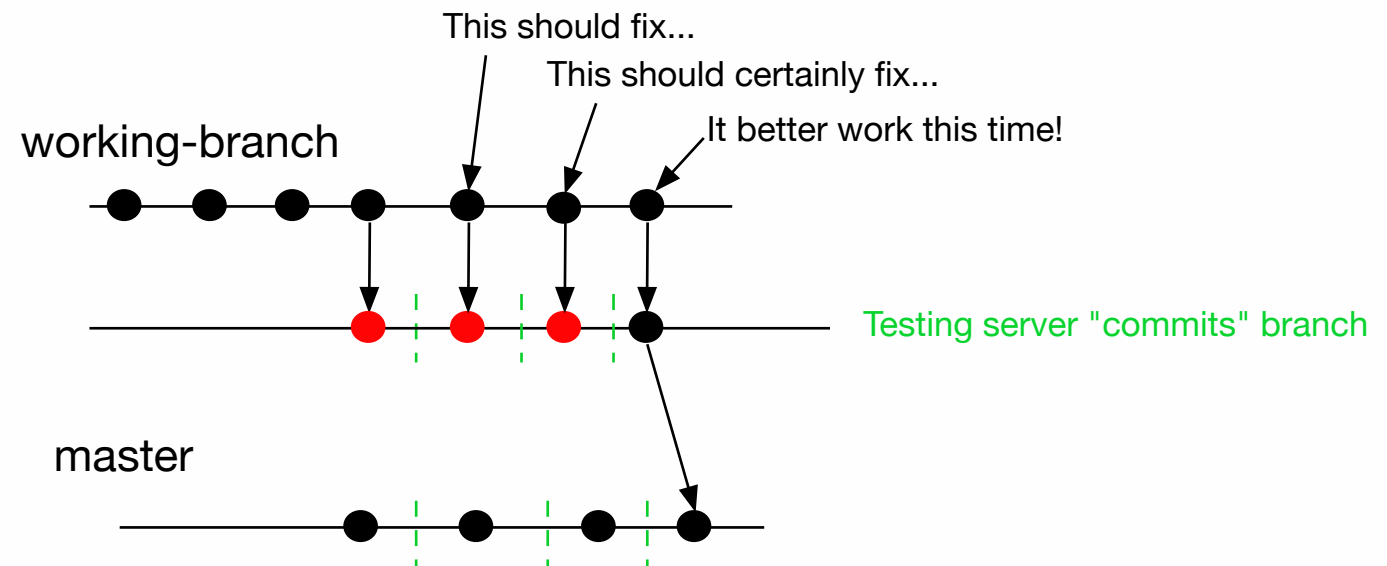
Old workflow:

working-branch

master

This should fix...

This should certainly fix...

Testing server runs

Pro-active testing workflow

Old workflow:

working-branch

This should fix...

This should certainly fix...

master

Testing server runs

New workflow:

This should fix...

This should certainly fix...

It better work this time!

working-branch

Testing server "commits" branch

master

Demo

http://benchmark.graylab.jhu.edu

You can submit a single revision for testing before you merge it to the master branch

- Commit your changes to non-master branch

- Push changes to GitHub

- Open Submit page http://benchmark.graylab.jhu.edu/submit enter commit sha1 and select test you want to run

- Instead of revision sha1 you can also submit:

  - branch name

  - pull-request-id

Special tests and branch selector on Submit page

Integration with GitHub Pull-Request system

Integration with GitHub Pull-Request system

● **Waiting to hear about a2ea624** — Queued for testing · Details

Merge with caution!
You can also merge branches on the command line.

🖥 | ⑂ **Merge pull request**

○— 🖼 CCD Refactor: few more unit tests —- (as an aside: what the heck is c… ···   ✖ e3ae979
○— 🖼 typo fix   ✖ 07800d6   **Assignee**
○— 🖼 merging in last good master state before pushing the BGB   ✔ 844bdae   🖼 **JWLabonte**

Success: Tests passed with regression tests difference. MERGE WITH CAUTION!

Integration with GitHub Pull-Request system

✖ Failed — Some build/unit tests got broken. DO NOT MERGE! · Details

Merge with caution!
You can also merge branches on the command line.

[🖥] [⑂ Merge pull request]

✔ All is well — Tests passed with regression tests difference. MERGE WITH CAUTION! · Details

This pull request can be automatically merged.
You can also merge branches on the command line.

[🖥] [⑂ Merge pull request]

Observers framework for Regression tests

- Add 'observers' text file to your integration tests dir

- Add lines with text: `branch1:email1 branch2:email2` ... for each branch you want to get for

- Example:
  `master:`<u>`sarel@weizmann.ac.il`</u>
  `master:`<u>`glapidoth@gmail.com`</u>

- De-couple observer info specification from TestServer code

- Allow to specify multiple observers for multiple branches

- If score/integration tests get renamed or copied observers info is preserved

RosettaLog@gmail.com <rosettalog@gmail.com>   Jul 22 (1 day ago)
to sarel, glapidoth

Test(s) mac.clang.integration.splice_out, linux.gcc.integration.debug.splice_out got broken by following commit:

sarel commited 2b97188dd4adf678 [diff] to branch master:57095 with commit message:

adding two integration tests for splice_in and splice_out for non-antibody cases

Tracking Custom Branches on Test Server:

- Testing server now tracks 5 branches

- We can customize which tests to run for each branch

- Could be *very* useful for a team of developers who work on the same project

Branch: master ▾   < 1 2 3 4 5 6 7 8 9 10 .. 42 >

benchmark
master
PyRosetta
release
rfalford12/mpframework_devel

er-Fay
aa3a7cd8f39」
ted commit: 🔍 code diff
4:59:00

	linux.clang	linux.gcc	mac.clang
debug	🟩	🟥	🟩
release	🟩	🟥	🟩
unit	🟩	🟥	🟩

Fixing signed/unsigned integer comparison (that I guess gcc doesn't catch?)

- Julia Koehler and Rebecca Alford (rfalford12/mpframework_devel) wrote:

  Benchmark-2 has been a really great resource for developing the membrane code. Frequently running all of the build, unit and integration tests helps us to keep our development compatible with master. Benchmark-2 is also a great collaboration tool - it helps Julia and I keep a running log of changes, and more easily merge our contributions.

News since last summer:

Daemon	Status	Test Running	Message	Time
Fujii	[ALIVE]	「master:57583」 mac.gcc.PyRosetta.unit	Running test PyRosetta.unit...	2015-02-03 12:40
Glass	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 13:59
Haruka	[ALIVE]	「master:57583」 windows.cl.PyRosetta.build	Running test PyRosetta.build...	2015-02-03 11:51
Hikaru	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Hojo-1	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Hojo-2	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Revision	[ALIVE]			2015-02-03 14:02
Rosetta	[ALIVE]	「master:57583」 linux.gcc.static.performance	Running test performance...	2015-02-03 13:33
Takeshi	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02

News since last summer:

Daemon	Status	Test Running	Message	Time
Fujii	[ALIVE]	「master:57583」 mac.gcc.PyRosetta.unit	Running test PyRosetta.unit...	2015-02-03 12:40
Glass	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 13:59
Haruka	[ALIVE]	「master:57583」 windows.cl.PyRosetta.build	Running test PyRosetta.build...	2015-02-03 11:51
Hikaru	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Hojo-1	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Hojo-2	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Revision	[ALIVE]			2015-02-03 14:02
Rosetta	[ALIVE]	「master:57583」 linux.gcc.static.performance	Running test performance...	2015-02-03 13:33
Takeshi	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02

- 9 machines now work on tests (was ~3 just this summer!)

News since last summer:

Daemon	Status	Test Running	Message	Time
Fujii	[ALIVE]	「master:57583」 mac.gcc.PyRosetta.unit	Running test PyRosetta.unit...	2015-02-03 12:40
Glass	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 13:59
Haruka	[ALIVE]	「master:57583」 windows.cl.PyRosetta.build	Running test PyRosetta.build...	2015-02-03 11:51
Hikaru	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Hojo-1	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Hojo-2	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Revision	[ALIVE]			2015-02-03 14:02
Rosetta	[ALIVE]	「master:57583」 linux.gcc.static.performance	Running test performance...	2015-02-03 13:33
Takeshi	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02

- 9 machines now work on tests (was ~3 just this summer!)

- We upgraded compilers to GCC-4.8, Clang-3.4 and created the protocol of adding new machines to the pool.

News since last summer:

Daemon	Status	Test Running	Message	Time
Fujii	[ALIVE]	「master:57583」 mac.gcc.PyRosetta.unit	Running test PyRosetta.unit...	2015-02-03 12:40
Glass	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 13:59
Haruka	[ALIVE]	「master:57583」 windows.cl.PyRosetta.build	Running test PyRosetta.build...	2015-02-03 11:51
Hikaru	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Hojo-1	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Hojo-2	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Revision	[ALIVE]			2015-02-03 14:02
Rosetta	[ALIVE]	「master:57583」 linux.gcc.static.performance	Running test performance...	2015-02-03 13:33
Takeshi	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02

- 9 machines now work on tests (was ~3 just this summer!)

- We upgraded compilers to GCC-4.8, Clang-3.4 and created the protocol of adding new machines to the pool.

- We finally finished our cluster upgrade and now i work on moving cluster tests to Benchmark-2 platform

News since last summer:

Daemon	Status	Test Running	Message	Time
Fujii	[ALIVE]	「master:57583」 mac.gcc.PyRosetta.unit	Running test PyRosetta.unit...	2015-02-03 12:40
Glass	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 13:59
Haruka	[ALIVE]	「master:57583」 windows.cl.PyRosetta.build	Running test PyRosetta.build...	2015-02-03 11:51
Hikaru	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Hojo-1	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Hojo-2	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02
Revision	[ALIVE]			2015-02-03 14:02
Rosetta	[ALIVE]	「master:57583」 linux.gcc.static.performance	Running test performance...	2015-02-03 13:33
Takeshi	[ALIVE]		Waiting for a next revision to be commited...	2015-02-03 14:02

- 9 machines now work on tests (was ~3 just this summer!)

- We upgraded compilers to GCC-4.8, Clang-3.4 and created the protocol of adding new machines to the pool.

- We finally finished our cluster upgrade and now i work on moving cluster tests to Benchmark-2 platform

- New release system. Rosetta-source, Rosetta-binaries and PyRosetta releases are available as tar.bz2 archives and Git repositories. And we can build release packages for custom commits on user request.

Queue Page http://benchmark.graylab.jhu.edu/queue

Daemon	Status	Test Running	Message	Time
Fujii	[ALIVE]	「benchmark:102」 mac.gcc.PyRosetta.unit	Running test PyRosetta.unit...	2015-02-04 14:55
Glass	[ALIVE]		Waiting for a next revision to be commited...	2015-02-04 15:23
Haruka	[ALIVE]	「master:57584」 windows.cl.PyRosetta.build	Running test PyRosetta.build...	2015-02-04 15:03
Hikaru	[ALIVE]	「PyRosetta:96」 mac.gcc.PyRosetta.unit	Running test PyRosetta.unit...	2015-02-04 14:46
Hojo-1	[ALIVE]	「benchmark:102」 linux.gcc.mpi.build.debug	Running test build.debug...	2015-02-04 15:22
Hojo-2	[ALIVE]	「PyRosetta:97」 linux.gcc.PyRosetta.unit	Running test PyRosetta.unit...	2015-02-04 14:35
Revision	[ALIVE]		Scheduling tests for revision master:57584	2015-02-04 15:24
Rosetta	[ALIVE]	「benchmark:103」 linux.gcc.build.release	Running test build.release...	2015-02-04 14:43
Takeshi	[ALIVE]	「PyRosetta:96」 linux.gcc.PyRosetta.unit	Running test PyRosetta.unit...	2015-02-04 14:34

Detailed Daemons Info 「click to collapse/expand」

Current Test Queue

Priority	Revision		Test	Queued Time
2000	「benchmark:102」	by Julia Koehler Leman	linux.gcc.static.build.release	2015-02-04 13:17
2000	「benchmark:102」	by Julia Koehler Leman	linux.gcc.static.performance	2015-02-04 13:17
2000	「benchmark:102」	by Julia Koehler Leman	linux.clang.mysql.build.debug	2015-02-04 13:17
2000	「benchmark:102」	by Julia Koehler Leman	linux.clang.postgres.build.debug	2015-02-04 13:17
2000	「benchmark:102」	by Julia Koehler Leman	mac.clang.opencl.build.debug	2015-02-04 13:17

It is now possible to develop and debug new tests locally (including cluster tests)

- 'benchmark.py' script emulates testing server API

```
% cd main/tests/benchmark
% ./benchmark.py -jN release.source
```

- See `main/tests/benchmark/tests/*` for code examples. Including the implementation of 'release' scripts and tests that use HPC (performance benchmark) runs.

Topics for discussion:

- What should we do with Performance benchmark false alarms?

- Revert policy: Should we revert pushes to master that broke things? (should this happened 'automatically' without consulting the author?)

Future Directions

Future Directions

- New dedicated machine for Window.PyRosetta builds is ordered and expected to arrive by the end of Feb 2015

Future Directions

- New dedicated machine for Window.PyRosetta builds is ordered and expected to arrive by the end of Feb 2015

- First version of HPC driver is now in master (and "Performance" test already uses it). Next steps:

Future Directions

- New dedicated machine for Window.PyRosetta builds is ordered and expected to arrive by the end of Feb 2015

- First version of HPC driver is now in master (and "Performance" test already uses it). Next steps:

  - Move all our Scientific tests to Benchmark-2 and refactor them along the way.

Future Directions

- New dedicated machine for Window.PyRosetta builds is ordered and expected to arrive by the end of Feb 2015

- First version of HPC driver is now in master (and "Performance" test already uses it). Next steps:

    - Move all our Scientific tests to Benchmark-2 and refactor them along the way.

    - Move Profile tests to Benchmark-2

Future Directions

- New dedicated machine for Window.PyRosetta builds is ordered and expected to arrive by the end of Feb 2015

- First version of HPC driver is now in master (and "Performance" test already uses it). Next steps:

  - Move all our Scientific tests to Benchmark-2 and refactor them along the way.

  - Move Profile tests to Benchmark-2

- Implement binary search for best-effort tests

Your feedback and suggestions are welcome!

Thank you!

http://benchmark.graylab.jhu.edu/submit

Submit Custom Revision for Testing

`65520983418baf5` ▶

Commited by: 「Labonte」
GitHub commit link: 「65520983418baf54c31541a61837c364d777345f」
Commit date: 2014-07-22T21:31:04Z

```
ResdidueProperties: improvements to auto-code-generation script

update_residue_properties.py now...
...only re-writes the files if changes have occured,
...ignores duplicated properties, and
...is more modular.


------------------------------------------
Unit test status: Not Run
Integration test status: Not Run
```

Rosetta Build Tests

☐ linux.clang.build.debug   ☐ linux.clang.build.release
☐ linux.gcc.build.debug   ☐ linux.gcc.build.release
☐ mac.clang.build.debug   ☐ mac.clang.build.release
☐ static

[none] [all]

Code Integrety Tests

☐ header only   ☐ library levels

[none] [all]

Unit Tests

☐ linux.clang.unit   ☐ linux.gcc.unit
☐ mac.clang.unit

[none] [all]

Regression Tests

☐ integration   ☐ score

[none] [all]

PyRosetta Tests

☐ linux.PyRosetta   ☐ mac.PyRosetta
☐ windows.PyRosetta
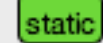
[none] [all]

Other Tests

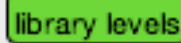☐ mpi   ☐ mysql
☐ postgres

[none] [all]

Tuesday, February 10, 15

Now we can comment on revision:

branch: _commits_ 「№10」
Commited by: Labonte
GitHub commit link: 「912b2aef1a6c1606」

Commit date: 2014-07-17 01:53:29

	linux.clang	linux.gcc	mac.clang
debug	🟩	🟩	🟩
release	🟩	🟩	🟩
unit	🟥	🟥	🟥

mac.PyRosetta mac.gcc.build.PyRosetta windows.PyRos

mpi header only library levels static integration sco

How the hell did that entire file get wiped?

Labonte 1 week
Sergey, I don't think the new integration tests are compared to the ref from when I branched from master.

Rosetta Benchmark 0 seconds
They compared to latest successful run from the master (i.e. run where build did not fail)

Leave a comment...

- So we can 'add' info to accidental commit messages

- And we can comment on test results and such