Testing Infrastructure in Rosetta

Sergey Lyskov

Jeff Gray lab JHU
2009

http://rosettatests.graylab.jhu.edu

Old data flow
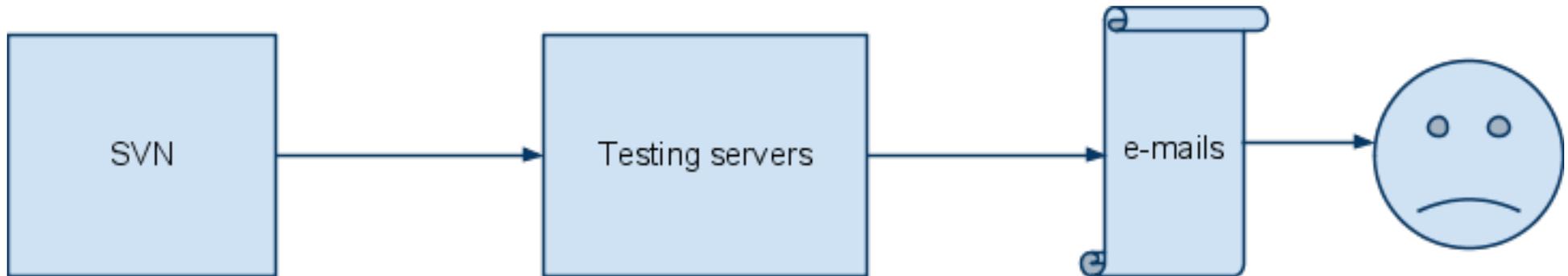
- A lot of email that no one wants to read!

- Hard to search for old results.

- Hard to compare old with new results.

- We can not send all data (logs/graphs/files) that user may need...

- Testing servers can not exchange data with each other, making tasks like generating summaries/tests report hard to implement.

- Realistically results can only consist of text logs...

New data flow
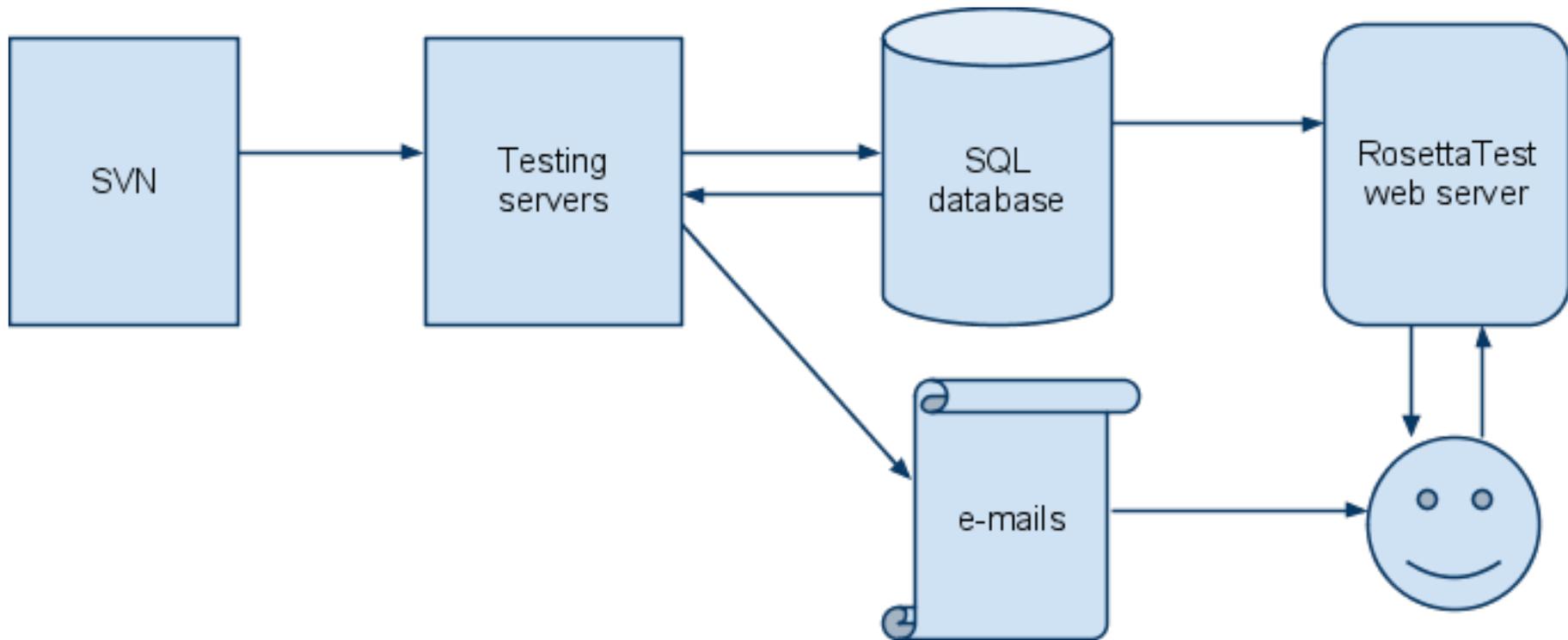
- We don't have to send so many large emails now, user could retrieve information on a need-to-know basis!
- All test results are searchable now!
- We can store full results of tests in DB (including all files) and provide it to user on request.
- Results could consist of logs/graphs/picture/pdf... anything!
- Side-by-side comparison of results of two revisions.
- Testing servers can retrieve previous results from DB and use it to compare data and generate summaries.

Available test types

- Unit tests

- Integration (regression) tests

- Performance benchmark tests

- Scientific daily tests (single CPU)

- Scientific weekly tests (single CPU)

- Scientific cluster tests (run on condor guided 80CPU cluster)

524 Unit tests, 41 Integration tests, 12 Scientific tests, 9 Performance tests

List of currently implemented scientific tests:

1. abinitio
2. dna_interface_design   **[provide pass/fail flag]**
3. docking
4. ligand_docking
5. loop
6. membrane
7. monomer_ddg
8. multi_residue_ligand_docking
9. relax
10. rna_denovo   **[provide pass/fail flag]**
11. rna_design   **[provide pass/fail flag]**
12. sequence_recovery

Test **'ScientificDaily.rna_design'** for SVN revison [31755]

Trac link: 31755

Test id: 9039   <-- send this for tech support
Test name: ScientificDaily.rna_design
Log file: [log]
YAML file: [yaml]
Test files: [files]
Started: 2009-07-29 19:47
Finished: 2009-07-29 22:24

You can view results of two revisions side by side by selecting second revisior
Please note that some revision may not contain YAML information or contain

Variable	31755	Plot
RNA_SeqRecoveryNoBasePairs	0.38	P
RNA_SeqRecoveryNonWC_BasePairs	0.74	P
RNA_SeqRecoveryOverall	0.52	P
RNA_SeqRecoveryWC_BasePairs	0.45	P
_isTestPassed	True	P

Test Log [Original unformated log] ---------------------------------

             no base pairs, seq. recovery: 0.38   [should be more than   0.30]
        Watson/Crick bp, seq. recovery: 0.45   [should be more than   0.35]
    non-Watson/Crick bp, seq. recovery: 0.74   [should be more than   0.70]
             all RNA, seq. recovery: 0.52   [should be more than   0.45]

RosettaTests

Plotting value **RNA_SeqRecoveryNoBasePairs** from test **ScientificDaily.rna_design**
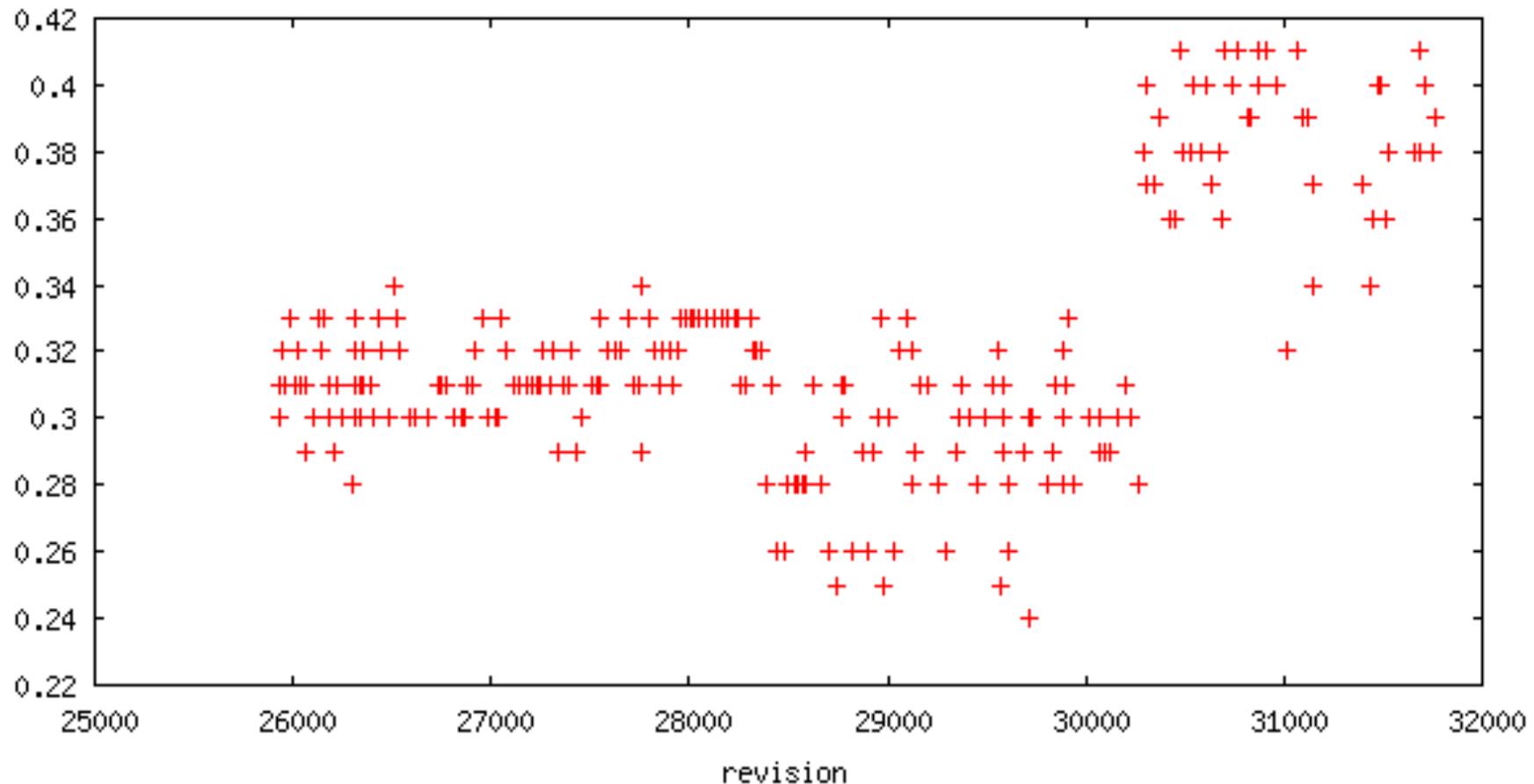
Available emails

- **[mini-logs]** 'Nightly tests results': Unit test + Integration tests + daily scientific tests results.

- **[mini-logs]** 'Daily change log': svn change log for the last 24h + build tests results.

- **[mini-logs]** 'Performance benchmark': Latest result of performance benchmark tests + comparison with previous results.

- **[rosetta++logs]** 'DailySummary': Provide daily overview and state of Mini branch which include build status, tests change/broken status, latest dates of successful run of scientific tests.

- **[personal]** 'Observer notification emails': Notify users if certain unit/integration test fails. (Each developer can customize this by providing list of tests that he want to 'observe')

- **[personal/mini-logs]** 'Commit confirmation email'/'Build/tests broken notification email': provide detailed information for each svn commit to main Mini trunk. This include: build status for each of 5 standard builds, results of unit tests, results of integration tests, C++ build warnings.

Testing Servers

- Revision-by-revision testing daemon. 8 CPU high-end Linux box. Run build/unit/integration for each revision.

- Nightly test daemon. Provide nightly SVN change log, daily scientific tests, performance benchmark tests.

- Weekly single CPU scientific test daemon.

- Cluster scientific test daemon. 104 CPU RosettaCommons dedicated cluster.

Fun facts

Since October 2008:

- ~ 9,000 tests suites ran

- ~ 900,000 individual tests executed.

- ~ 9,000 emails was sent to users (~250Mb).

http://rosettatests.graylab.jhu.edu/stats

SVN users statistics

User	Revisions commited	Builds broken (%)	Unit tests broken (%)	Integration tests changed (%)
tex	301	31 (10%)	15 (4%)	57 (18%)
yab	236	3 (1%)	5 (2%)	26 (11%)
mtyka	205	10 (4%)	3 (1%)	66 (32%)
sarel	194	4 (2%)	1 (0%)	35 (18%)
jecorn	184	7 (3%)	2 (1%)	28 (15%)
smlewis	143	2 (1%)	0 (0%)	24 (16%)

Brief overview of RosettaTests web site features.

http://rosettatests.graylab.jhu.edu

RosettaTests

List of relevant SVN revisions:

[all] [broken builds] [broken unit tests] [integration tests diff]

Revision	Trac link	User	Message	Status	Date
31753	T	sarel	adding PlaceSimultaneouslyMover. This should complement PlaceStubMover when placing more than just two hotspots. Very briefly, the idea is to triage very early if not all hotspots can be incorporated by the scaffold and then concentrate on perfecting the placements using bb motions. This is still work in progress. No test changes expected.	I	2009-07-29 00:07:35
31749	T	ekellogg		Ok	2009-07-28 22:24:00
31748	T	ekellogg		B	2009-07-28 22:23:06
31747	T	blivens	Fixed a bug which broke the graphics build.	B	2009-07-28 21:24:03
31746	T	dimaio	small error fix to symmfile generation script	B	2009-07-28 20:44:55
31745	T	ekellogg	added options for compatibility with minirosetta application	B	2009-07-28 20:44:05

RosettaTests

Finished Tests list.

Use this page to check which of the test suites was run for a given revision. [If you are looking for an individual Build, Unit or Integration test results, then use this page]

Starting revision: [ ] Ending revision: [ ] Category: [all ▼] Test name: [ ] [Filter]

ID	Revision	Test	Started	Finished
9006	31754	Build	2009-07-29 12:35	2009-07-29 13:11
9007	31754	UnitTests	2009-07-29 13:11	2009-07-29 13:25
9008	31754	IntegrationTests	2009-07-29 13:25	2009-07-29 13:30
9002	31753	Build	2009-07-29 03:45	2009-07-29 04:21
9003	31753	UnitTests	2009-07-29 04:21	2009-07-29 04:36
9004	31753	IntegrationTests	2009-07-29 04:36	2009-07-29 04:40
9005	31753	ScientificWeekly.monomer_ddg	2009-07-29 08:18	2009-07-29 12:03
8999	31749	Build	2009-07-29 02:49	2009-07-29 03:25
9000	31749	UnitTests	2009-07-29 03:25	2009-07-29 03:39
9001	31749	IntegrationTests	2009-07-29 03:39	2009-07-29 03:44
8998	31748	Build	2009-07-29 02:14	2009-07-29 02:48
8997	31747	Build	2009-07-29 01:39	2009-07-29 02:13
8996	31746	Build	2009-07-29 01:04	2009-07-29 01:38
8995	31745	Build	2009-07-29 00:29	2009-07-29 01:03
8994	31744	Build	2009-07-28 23:54	2009-07-29 00:28
8991	31743	Build	2009-07-28 22:58	2009-07-28 23:34
8992	31743	UnitTests	2009-07-28 23:34	2009-07-28 23:48
8993	31743	IntegrationTests	2009-07-28 23:48	2009-07-28 23:52

RosettaTests

Search for individual Build, Unit or Integration tests:

Use this page to search for test results within following test suites: Build, Unit or Integration. [If you are looking for a particular scientific or performance test then use this page]

Starting revision: [　　] Ending revision: [31740] Test name: [　　] Status: [failed ▾] [Search]

ID	Test ID	Revision	Test	Status
953162	8979	31733	Integration.enzdes	Failed
952596	8976	31729	Integration.metalloprotein_broker	Failed
952582	8976	31729	Integration.broker	Failed
952565	8976	31729	Integration.ThreadingInputter	Failed
948043	8943	31681	Build.pilot_apps_all	Failed
948042	8943	31681	Build.unit_tests	Failed
948041	8943	31681	Build.pilot_apps_all_release	Failed
948040	8943	31681	Build.libraries_release	Failed
948039	8943	31681	Build.libraries	Failed
948038	8942	31680	Build.pilot_apps_all	Failed
948037	8942	31680	Build.unit_tests	Failed
948036	8942	31680	Build.pilot_apps_all_release	Failed
948035	8942	31680	Build.libraries_release	Failed
948034	8942	31680	Build.libraries	Failed

RosettaTests

Latest tested revision is: [31754]

```
Test Log [Original unformated log] ------------------------------
BUILDS BROKEN:                  0 (out of 5)
COMPILER WARNINGS:              8
UNIT TESTS BROKEN:              0 (out of 524)
INTEGRATION TESTS DIFFERENT: 0 (out of 41)


Message: correct comments
Build status: -------------------------------------
  libraries                - Ok.
  libraries_release        - Ok.
  pilot_apps_all           - Ok.
  pilot_apps_all_release - Ok.
  unit_tests               - Ok.


Integration tests status: --------------------------

Build warinings: -----------------------------------
dekim:30434
  src/protocols/moves/TailsScoreMover.cc:108: warning: converting to `int' from
`core::Real'
  src/protocols/moves/TailsScoreMover.cc:126: warning: converting to `int' from
`core::Real'

ekellogg:31682
  src/protocols/ddg/ddG_main.cc:95: warning: unused parameter
'mut_avg_components'

ekellogg:31749
  src/protocols/moves/ddGMover.cc:1031: warning: comparison between signed and
unsigned integer expressions
  src/protocols/moves/ddGMover.cc:1092: warning: unused variable 'debug_unbound'
  src/protocols/moves/ddGMover.cc:1087: warning: unused variable 'debug_bound'
  src/protocols/moves/ddGMover.cc:887: warning: comparison between signed and
unsigned integer expressions
```

RosettaTests

Test '**Performance**' for SVN revison [31523]

```
Trac link:    31523

Test id:      8747  <-- send this for tech support
Test name:    Performance
Log file:     [log]
YAML file:    [yaml]
Test files:   [files]
Started:      2009-07-19 15:59
Finished:     2009-07-19 16:20
```

You can view results of two revisions side by side by selecting second revision number 30959 ▾ and click Compare
Please note that some revision may not contain YAML information or contain only partial YAML information.

Variable	31523	30959	Plot
core_scoring_Score	81.119668	80.64874	P
protocols_docking_DockingHighRes	171.830878	174.50947	P
protocols_docking_DockingLowRes	207.060522	211.314875	P
protocols_ligand_docking_LigandDockProtocol	40.351866	40.968772	P
protocols_moves_ShearMover	81.692581	81.741574	P
protocols_moves_SmallMover	84.893094	87.347721	P
protocols_optimization_Minimizer_dfpmin	133.300736	132.15291	P
protocols_optimization_Minimizer_dfpmin_armijo	149.349295	147.969505	P
protocols_optimization_Minimizer_dfpmin_armijo_nonmonotone	148.769384	150.734084	P

RosettaTests

Plotting value **core_scoring_Score** from test **Performance**
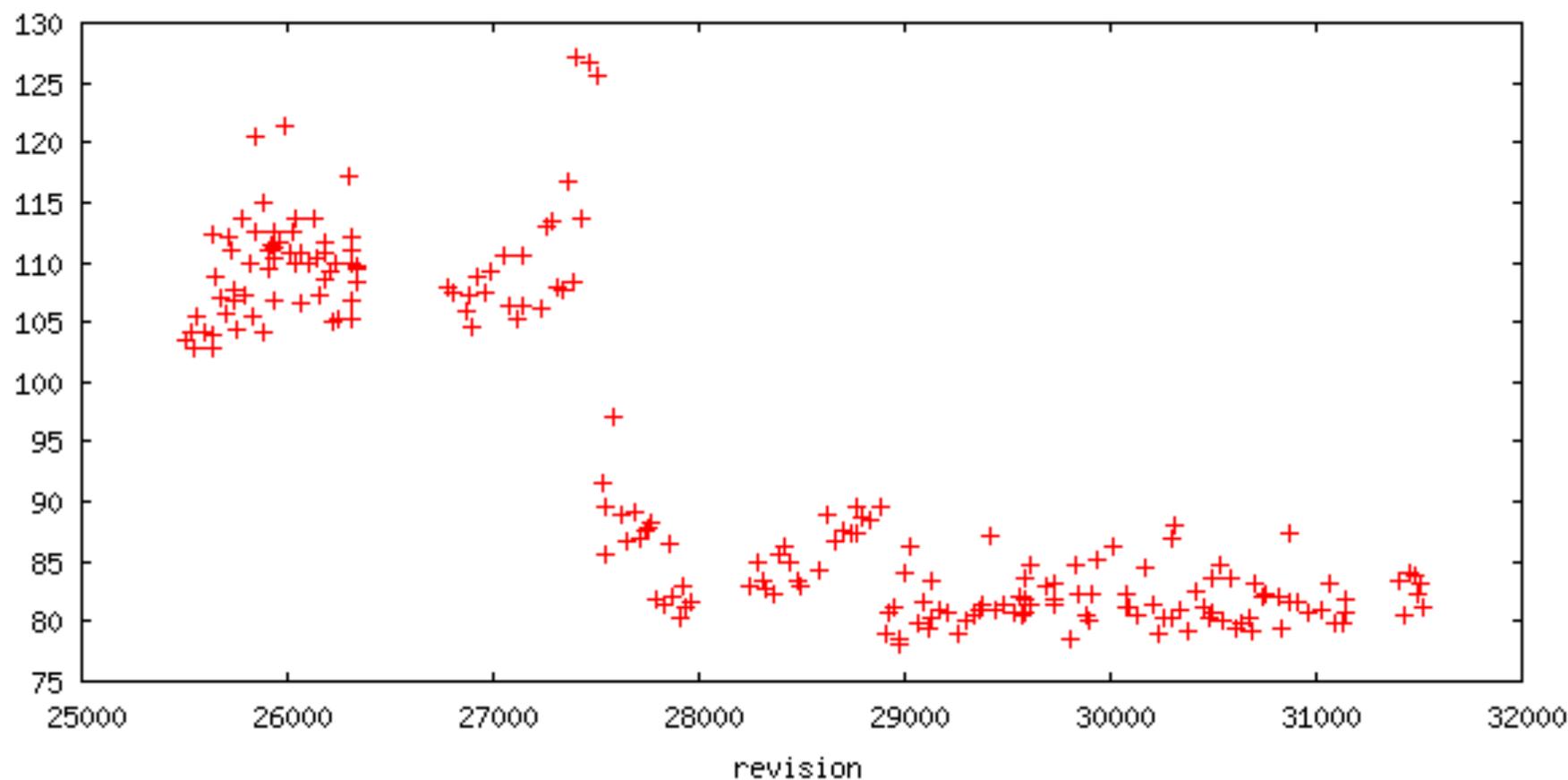
RosettaTests

This is predefined plots for most commonly used variables. You can create your own plot by selecting: tests, select test that contain variable you want to plot, and clicking on plot link ('P'). Note: Only variable that present in YAML output can be plotted (most scientific tests contain all important results in YAML form).

Performance benchmark

core_scoring_Score

protocols_docking_DockingHighRes

protocols_docking_DockingLowRes

protocols_moves_PackRotamersMover

protocols_moves_ShearMover

protocols_moves_SmallMover

protocols_optimization_Minimizer_dfpmin

protocols_optimization_Minimizer_dfpmin_armijo

protocols_optimization_Minimizer_dfpmin_armijo_nonmonotone

Sequence recovery

HydrophobicNon_PolarAminoAcids

IdentityForAllPositions

NegativeChargedAminoAcid

PolarUnchargedAmino

PositiveChargedAminoAcid

Next priorities

1. Assist users in adding pass/fail criteria to scientific benchmarks.

2. Add better performance (speed) benchmarks.

3. Upgrade to gcc 4.1!

4. Simplify debugging of the tests.

   o Create a way to debug cluster scientific tests on single CPU machines.

   o Add option to download testing server binaries.

5. Add more features to website to improve usability.

Suggestions and feedback welcome!