Update on
Testing Infrastructure in Rosetta

Sergey Lyskov

Jeff Gray lab JHU
2010

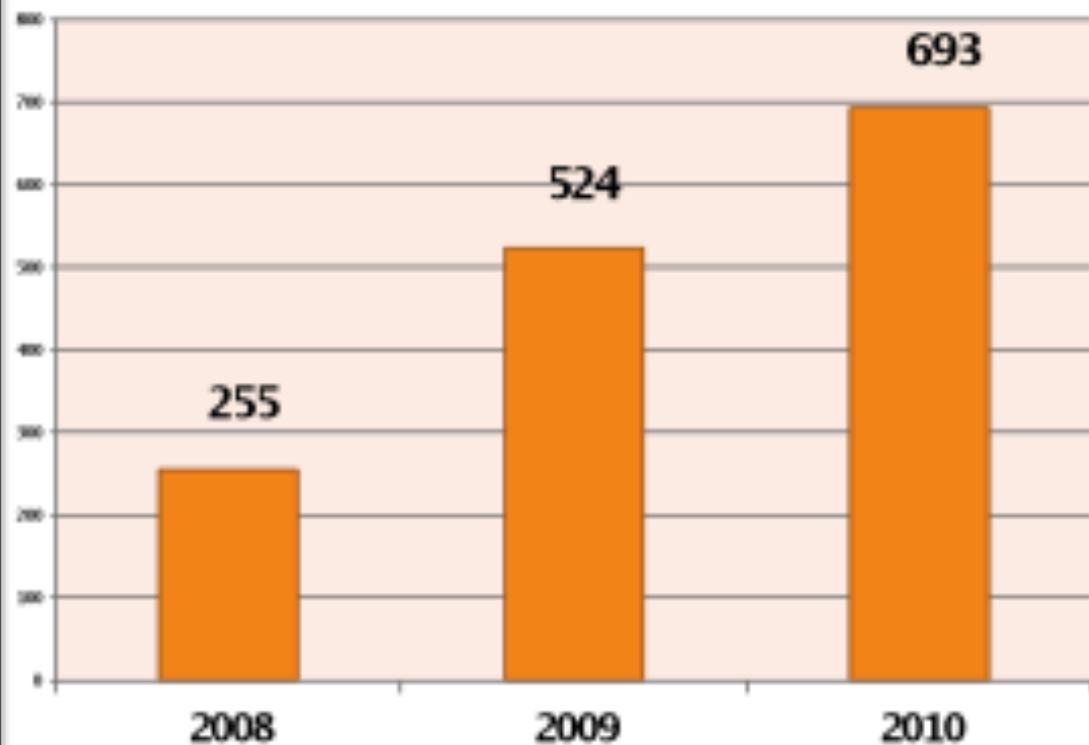http://rosettatests.graylab.jhu.edu
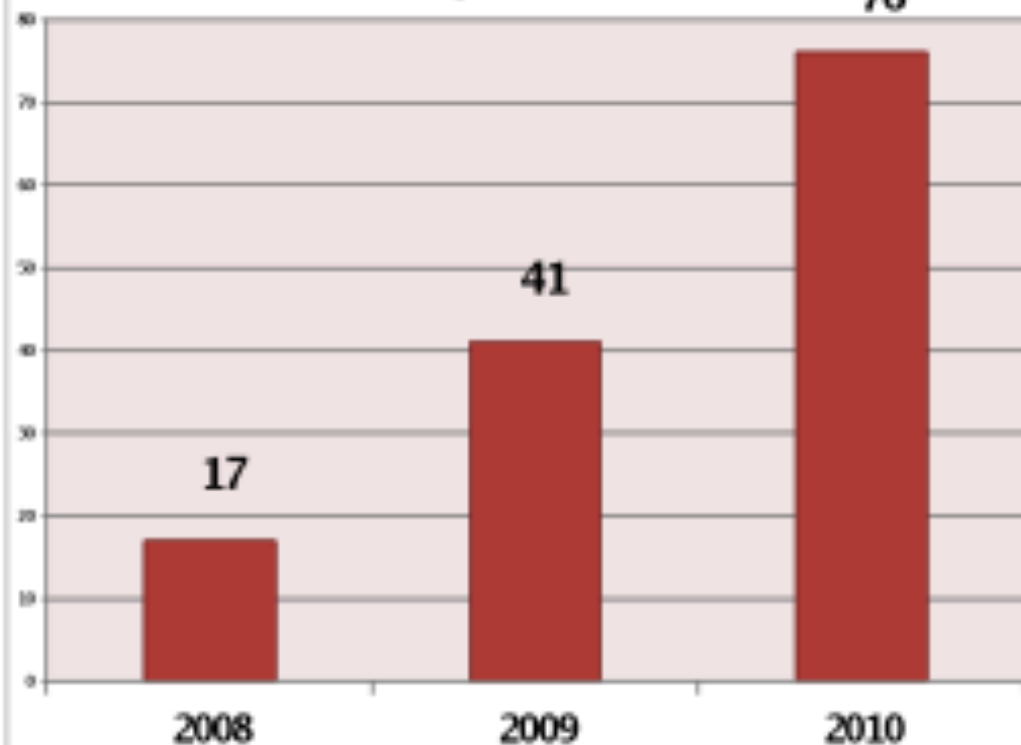
Quick overview for new members

Types of tests that we currently run:

- Build tests

- Unit tests

- Integration tests

- Performance tests

- Scientific tests

- Profile test
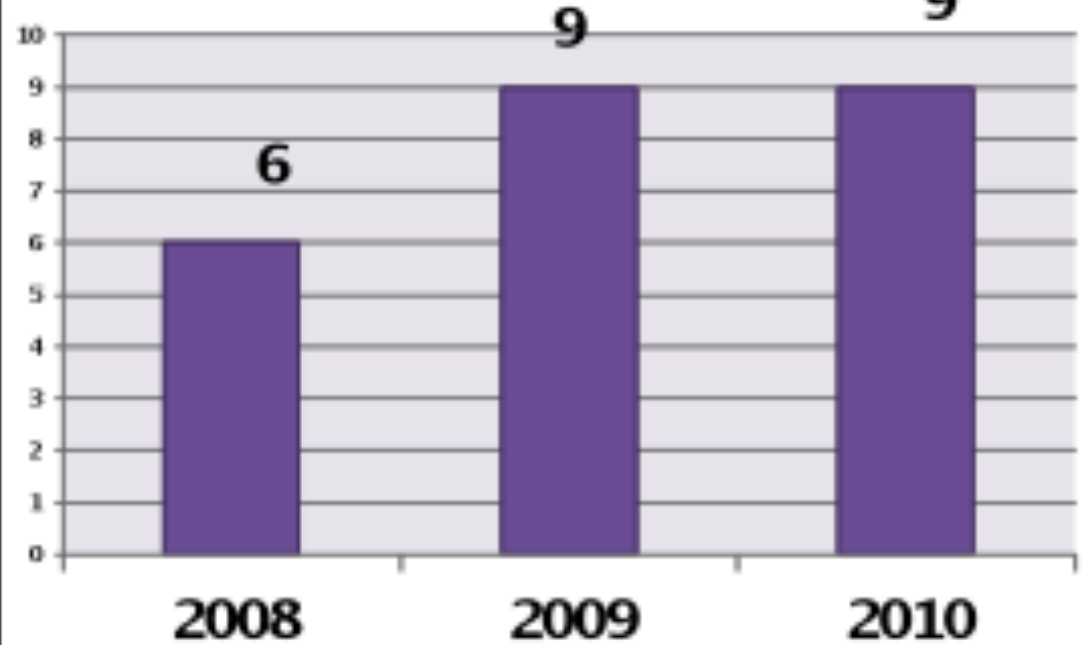
Unit tests

- 2008: 255
- 2009: 524
- 2010: 693

Integration tests

- 2008: 17
- 2009: 41
- 2010: 76

Performance tests

- 2008: 6
- 2009: 9
- 2010: 9

Scientific tests

- 2008: 3
- 2009: 12
- 2010: 15

Thursday, August 5, 2010
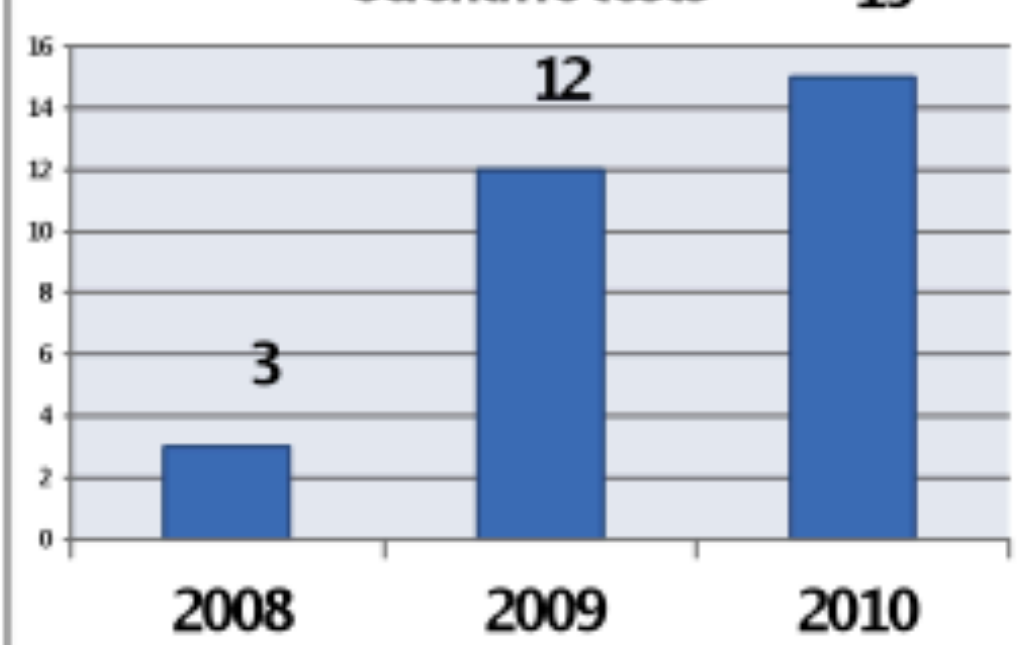
New test type: Profile Tests, currently **9** tests implemented.

AnchoredDesign memory profile:
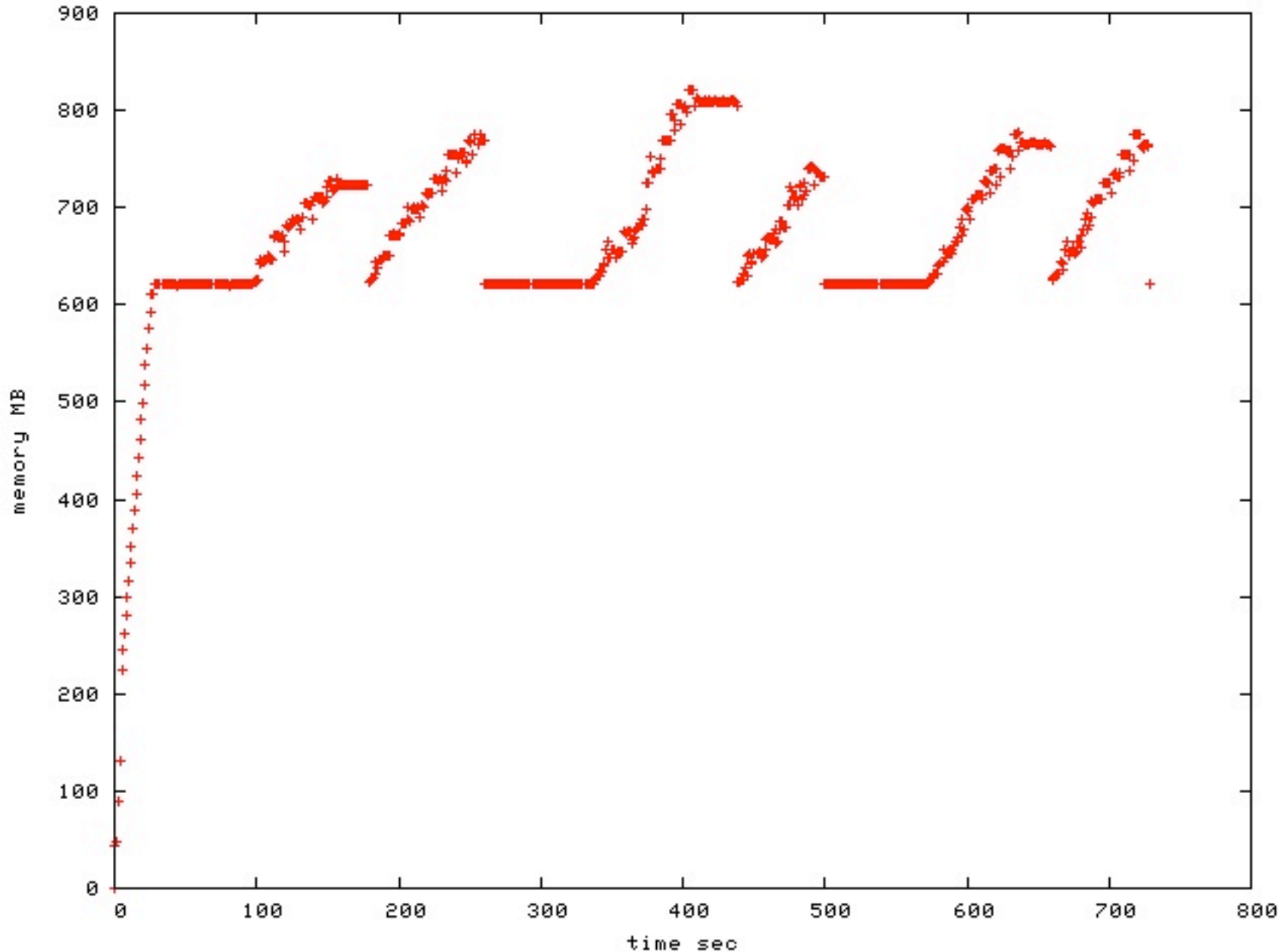
New build type: **HEADERS ONLY**

For this test for each header file we compile dummy C++ file that includes only our one header file.

•Helps keep include files in good shape.

•Helps maintain PyRosetta build.

- Now our test server automatically checks every mini revision for this build type and informs users if it got broken.

SVN Statistics!

Since last year:

- Revisions committed in to mini trunk: **2,187 ~ 6** rev./day

- Test suites run: **10,026**

- Number of individual tests executed: **~1,200,000**

Comparing Integration test files from two arbitrary revisions:

Files diff for tests IntegrationTests id=18853 [r37478] and id=1868

To view 'diff' output please select tests id/revisions you wish to compare:

18853 [r37478] ⬍   18689 [r37430] ⬍   Perform brief 'diff' ⬍   and click   Compare

Please note that some comparison requare to process almost 0.5Gb of data, so it could take a while.

```
Comparing tests [FullDiff=False]:
    IntegrationTests Test_id=18689 [r37430]
    IntegrationTests Test_id=18853 [r37478]


diff -rq 37478/ 37430/
Files 37478/antibody/aaFR02_0001.pdb and 37430/antibody/aaFR02_0001.pdb differ
Files 37478/antibody/aascore.fasc and 37430/antibody/aascore.fasc differ
Files 37478/antibody/flags and 37430/antibody/flags differ
Files 37478/antibody/log and 37430/antibody/log differ
Files 37478/assemble_domains_jd2/flags and 37430/assemble_domains_jd2/flags differ
Files 37478/assemble_domains_jd2/output/dock.out and 37430/assemble_domains_jd2/output/dock.out differ
Files 37478/assemble_domains_jd2/output/scores.fasc and 37430/assemble_domains_jd2/output/scores.fasc differ
Files 37478/cluster/log and 37430/cluster/log differ
Files 37478/enzdes/enz_score.out and 37430/enzdes/enz_score.out differ
```

Test files revisions tracking:
Summary grouped by Test

Test 'IntegrationTests' for SVN revison [37478]

Trac link: 37478

Test id: 18853   <-- send this for tech support
Test name: IntegrationTests
Log file: [log]
YAML file: [yaml]
Test files: [files] You can see 'diff' results for tests files from two different revisions here.
Started: 2010-07-19 22:00
Finished: 2010-07-19 22:05

IntegrationTest files have not changed since revision: [@37460] [T37460]

Individual tests output files history:

Test name	Same since revision
AnchoredDesign	[@37201] [T37201]
AnchoredPDBCreator	[@37193] [T37193]
FloppyTail	[@37193] [T37193]
HOW_TO_MAKE_TESTS	[@35263] [T35263]
RescorePDDF	[@37193] [T37193]
RescoreSAXS	[@37193] [T37193]
ThreadingInputter	[@37193] [T37193]
abinitio	[@37193] [T37193]
angle_recovery_stats	[@37193] [T37193]

Tracking history of individual test files:

[Revisions] [Tests] [Search] [CurrentStatus] [Analyze] [IntegrationTests Diff] [Statistics]

Files for test [IntegrationTests] at revision: [37478]

You can see 'diff' results for tests files from two different revisions here.

[@revision since file have not changed] file

[@35263] [T35263] jd2test/1UBQ.pdb.gz

[@35263] [T35263] jd2test/resfile

[@35263] [T35263] jd2test/options

[@35263] [T35263] jd2test/1EM7.pdb.gz

[@37193] [T37193] jd2test/command

[@35263] [T35263] jd2test/1NLO.pdb.gz

[@35263] [T35263] jd2test/1UBQ.pdb

[@35263] [T35263] jd2test/1EM7.pdb

[@35263] [T35263] jd2test/1NLO.pdb

[@37193] [T37193] jd2test/log

User who committed most revisions: **Sarel Fleishman**

User	Revisions commited	Builds broken (%)	Unit tests broken (%)	Integration tests changed (%)
sarel	204	7 (3%)	2 (0%)	47 (23%)
tex	201	12 (5%)	1 (0%)	53 (26%)
sergey	167	2 (1%)	1 (0%)	9 (5%)
smlewis	120	0 (0%)	2 (1%)	28 (23%)
jecorn	86	4 (4%)	0 (0%)	10 (11%)
dgront	85	11 (12%)	1 (1%)	27 (31%)
mtyka	82	3 (3%)	1 (1%)	22 (26%)
ekellogg	74	8 (10%)	0 (0%)	5 (6%)
leaverfa	71	5 (7%)	1 (1%)	15 (21%)
flo	68	1 (1%)	1 (1%)	13 (19%)
wendao	65	2 (3%)	0 (0%)	8 (12%)
olange	64	12 (18%)	0 (0%)	22 (34%)

User with highest rate of broken builds: **Oliver Lange**

Time frame: [Since last year ▲▼] Sort by: [Builds broken ▲▼] (Go!)

User	Revisions commited	Builds broken (%)	Unit tests broken (%)	Integration tests changed (%)
olange	64	12 (18%)	0 (0%)	22 (34%)
possu	36	5 (13%)	0 (0%)	7 (19%)
aroop	38	5 (13%)	0 (0%)	12 (31%)
dgront	85	11 (12%)	1 (1%)	27 (31%)
ekellogg	74	8 (10%)	0 (0%)	5 (6%)
barak	20	2 (10%)	0 (0%)	7 (35%)
sid	42	4 (9%)	2 (4%)	11 (26%)
renfrew	21	2 (9%)	1 (4%)	8 (38%)
sheffler	43	4 (9%)	0 (0%)	2 (4%)
rhiju	44	4 (9%)	4 (9%)	10 (22%)
andre	49	4 (8%)	0 (0%)	11 (22%)
glemmon	25	2 (8%)	0 (0%)	14 (56%)

User with highest rate of broken Unit tests: **Ron Jacak**

User	Revisions commited	Builds broken (%)	Unit tests broken (%)	Integration tests changed (%)
ronj	17	0 (0%)	2 (11%)	4 (23%)
rhiju	44	4 (9%)	4 (9%)	10 (22%)
krishna	14	1 (7%)	1 (7%)	6 (42%)
rvernon	36	1 (2%)	2 (5%)	7 (19%)
sid	42	4 (9%)	2 (4%)	11 (26%)
renfrew	21	2 (9%)	1 (4%)	8 (38%)
momeara	44	0 (0%)	2 (4%)	13 (29%)
dimaio	56	1 (1%)	2 (3%)	18 (32%)
smlewis	120	0 (0%)	2 (1%)	28 (23%)
flo	68	1 (1%)	1 (1%)	13 (19%)
leaverfa	71	5 (7%)	1 (1%)	15 (21%)
mtyka	82	3 (3%)	1 (1%)	22 (26%)

User with highest rate of Integration tests change: **Gordon Lemmon**

User	Revisions commited	Builds broken (%)	Unit tests broken (%)	Integration tests changed (%)
glemmon	25	2 (8%)	0 (0%)	14 (56%)
krishna	14	1 (7%)	1 (7%)	6 (42%)
achen39	17	0 (0%)	0 (0%)	7 (41%)
renfrew	21	2 (9%)	1 (4%)	8 (38%)
delucasl	11	0 (0%)	0 (0%)	4 (36%)
barak	20	2 (10%)	0 (0%)	7 (35%)
olange	64	12 (18%)	0 (0%)	22 (34%)
dimaio	56	1 (1%)	2 (3%)	18 (32%)
dgront	85	11 (12%)	1 (1%)	27 (31%)
aroop	38	5 (13%)	0 (0%)	12 (31%)
ashworth	38	2 (5%)	0 (0%)	12 (31%)
momeara	44	0 (0%)	2 (4%)	13 (29%)

Next priorities

1. Multi-Platform testing: Linux 32, Linux 64, Mac32, Mac64. Windows (cygwin)

2. Additional automatic builds testing: PyRosetta.

3. Possible cluster expansion (Do we need more nodes for scientific tests?)

4. We already moved our main testing server to GCC 4.1. Should we move to 64Bit environment as well?

Suggestions and feedback welcome!

Huge thanks to everyone who helped develop and maintain these tests!

List of currently implemented scientific tests:

1. abinitio
2. dna_interface_design
3. docking
4. ligand_docking
5. loop
6. membrane
7. monomer_ddg
8. multi_residue_ligand_docking
9. relax
10. rna_denovo
11. rna_design
12. sequence_recovery
13. detailed_balance
14. enzdes_benchmark
15. Rotamer Recovery

New things since last year

- Unit tests 524 → **693**.
- Integration tests 41 → **76**
- Performance test 9 → **9**
- Scientific tests 12 → **14**
- New Test type**: Profile tests → 9**

New build type: **Headers only**

New RosettaTests web site features:
- File changes tracking in a manner similar to SVN
- Generating Integration tests diff for two arbitrary revisions

We move our test server and now use gcc 4.1!