Finding a Good Length for Automatically Generated Tests

TABLE I
PYTHON SUBJECTS

SUT	LOC	Source
AVL	225	TSTL example [6]
bidict	569	http://pythonhosted.org/bidict/home.html
CParser	5033	https://github.com/albertz/PyCParser
pyfakefs	2642	https://github.com/jmcgeheeiv/pyfakefs
redis-py	2722	https://github.com/andymccurdy/redis-py
python-rsa	1597	https://github.com/sybrenstuvel/python-rsa
simplejson	2811	https://simplejson.readthedocs.io/en/latest/
sortedcontainers	2017	http://www.grantjenks.com/docs/sortedcontainers/
SymPy	227959	http://www.sympy.org/en/index.html

I. Introduction

There are a lot of papers out there on this topic [1]–[5].

II. Some Python Results

Table I describes the Python subjects, tested using the TSTL [7], [8] harnesses included in the TSTL distribution. Of the eight subjects, one is a toy container class with an injected realistic fault, and the other seven are real-world Python libraries, with a large number of GitHub stars or high ranking in pip. Results are shown for both standard random testing and swarm testing [9], using 120 seconds of test time. While the change to swarm testing sometimes changes the effectiveness curve, shapes at 30 seconds and 60 seconds are very similar, despite changes in overall coverage obtained.

III. Related Work

References

[1] J. H. Andrews, A. Groce, M. Weston, and R.-G. Xu, "Random test run length and effectiveness," in *Automated Software Engineering*, 2008, pp. 19–28.

[2] A. Arcuri, "Longer is better: On the role of test sequence length in software testing," in *Third International Conference on Software Testing, Verification and Validation, ICST 2010, Paris, France, April 7-9, 2010*. IEEE Computer Society, 2010, pp. 469–478. [Online]. Available: https://doi.org/10.1109/ICST.2010.16

[3] G. Fraser and A. Arcuri, "It is not the length that matters, it is how you control it," in *Fourth IEEE International Conference on Software Testing, Verification and Validation, ICST 2011, Berlin, Germany, March 21-25, 2011*. IEEE Computer Society, 2011, pp. 150–159. [Online]. Available: https://doi.org/10.1109/ICST.2011.54

[4] A. Arcuri, "A theoretical and empirical analysis of the role of test sequence length in software testing for structural coverage," *IEEE Trans. Software Eng.*, vol. 38, no. 3, pp. 497–519, 2012. [Online]. Available: https://doi.org/10.1109/TSE.2011.44

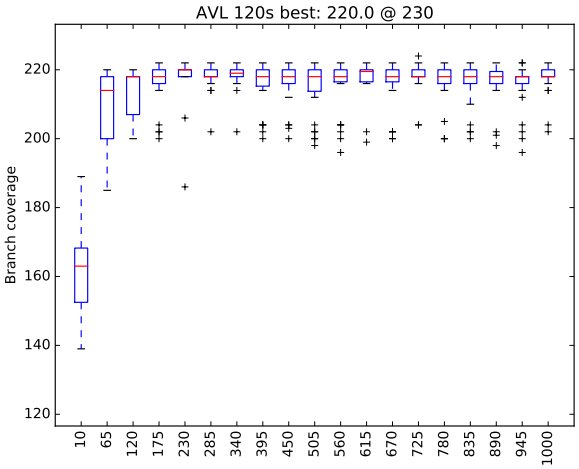
[5] G. Fraser and A. Arcuri, "Handling test length bloat," *Softw. Test., Verif. Reliab.*, vol. 23, no. 7, pp. 553–582, 2013. [Online]. Available: https://doi.org/10.1002/stvr.1495

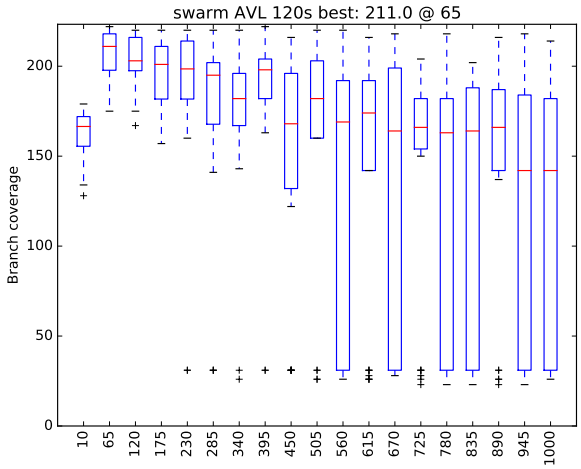
[6] user1689822, "python AVL tree insertion," http://stackoverflow.com/questions/12537986/python-avl-tree-insertion.

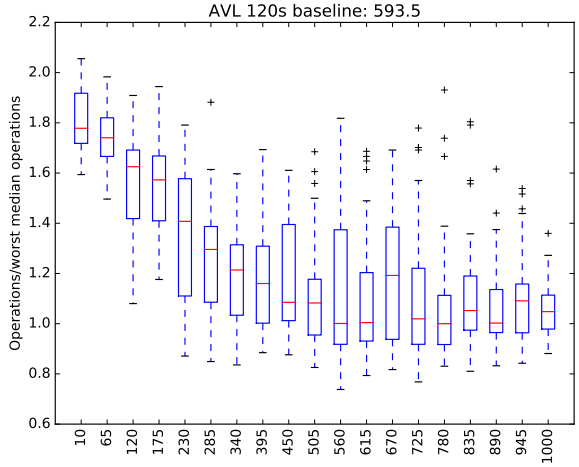
[7] J. Holmes, A. Groce, J. Pinto, P. Mittal, P. Azimi, K. Kellar, and J. O'Brien, "TSTL: the template scripting testing language," *International Journal on Software Tools for Technology Transfer*, 2017, accepted for publication.

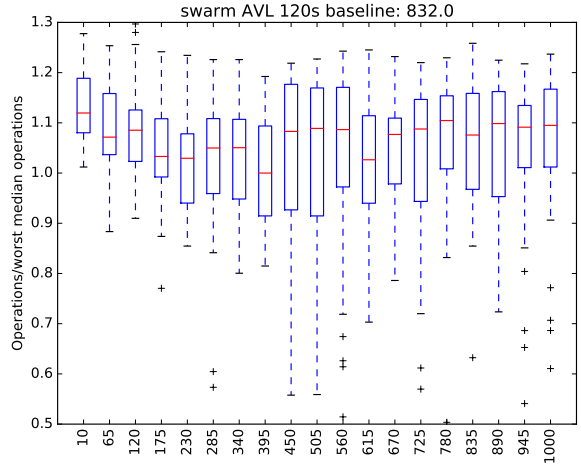
[8] A. Groce and J. Pinto, "A little language for testing," in *NASA Formal Methods Symposium*, 2015, pp. 204–218.

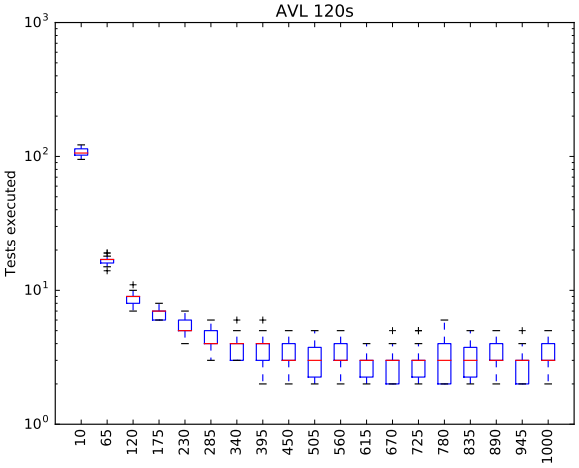
[9] A. Groce, C. Zhang, E. Eide, Y. Chen, and J. Regehr, "Swarm testing," in *International Symposium on Software Testing and Analysis*, 2012, pp. 78–88.

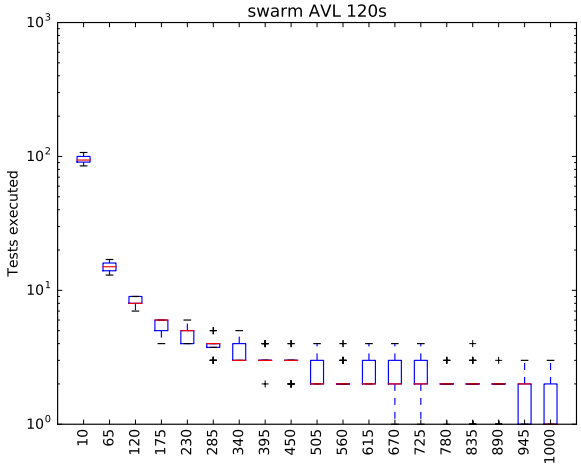

AVL 120s best: 220.0 @ 230
Branch coverage
120
140
160
180
200
220
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

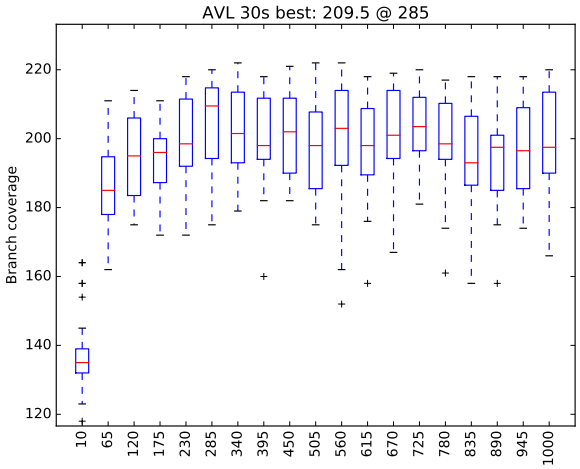

swarm AVL 120s best: 211.0 @ 65
Branch coverage
0
50
100
150
200
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

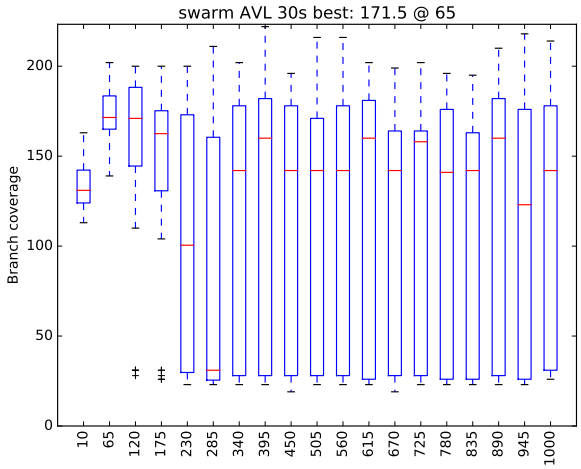

AVL 120s baseline: 593.5
Operations/worst median operations
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

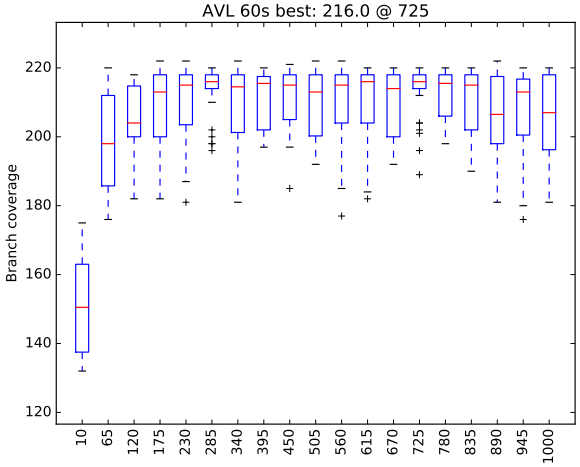

swarm AVL 120s baseline: 832.0
Operations/worst median operations
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

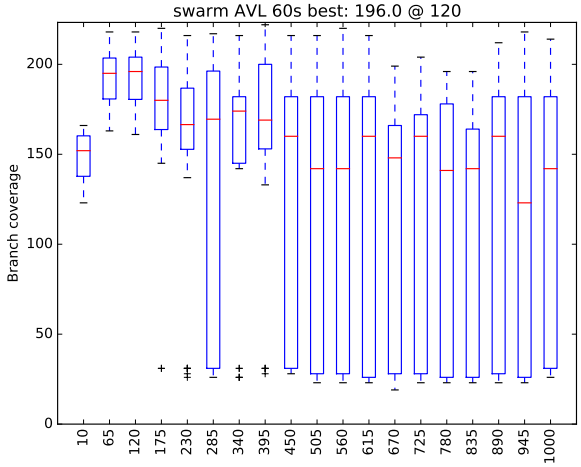

AVL 120s
Tests executed
10^0
10^1
10^2
10^3
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

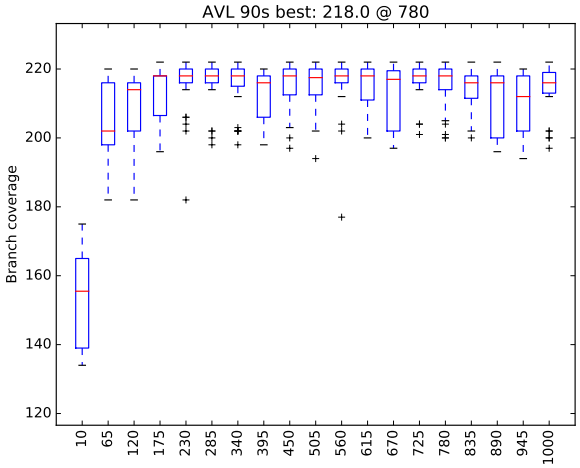

swarm AVL 120s
Tests executed
10^0
10^1
10^2
10^3
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

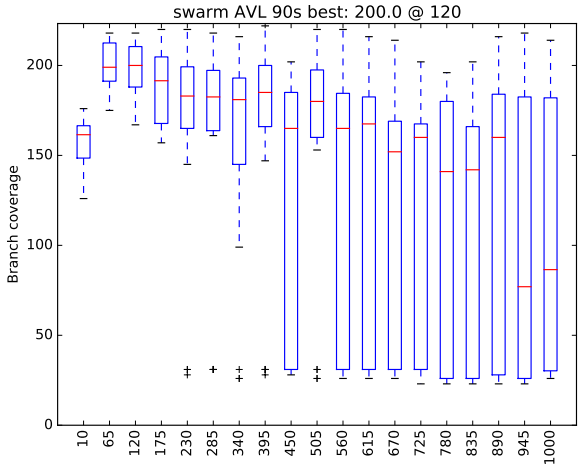

AVL 30s best: 209.5 @ 285
Branch coverage
120
140
160
180
200
220
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

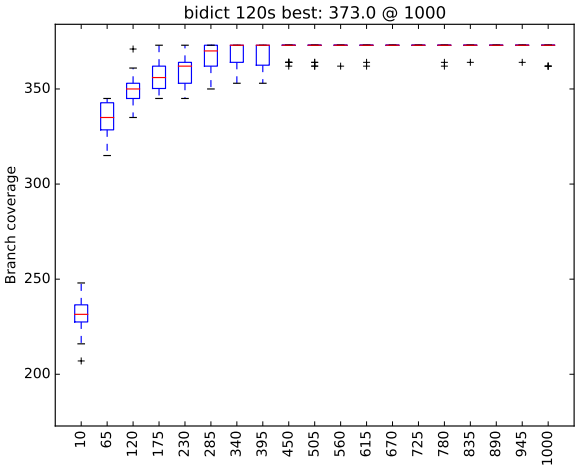

swarm AVL 30s best: 171.5 @ 65
Branch coverage
0
50
100
150
200
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

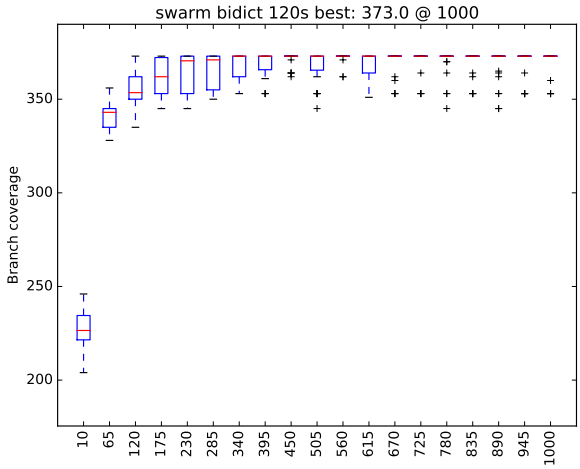

AVL 60s best: 216.0 @ 725
Branch coverage
120
140
160
180
200
220
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

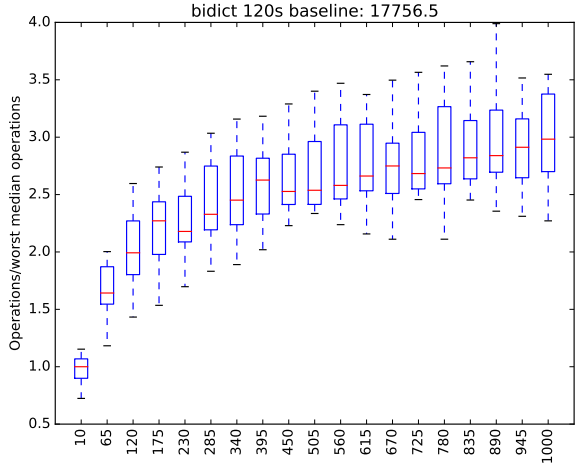

swarm AVL 60s best: 196.0 @ 120
Branch coverage
0
50
100
150
200
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

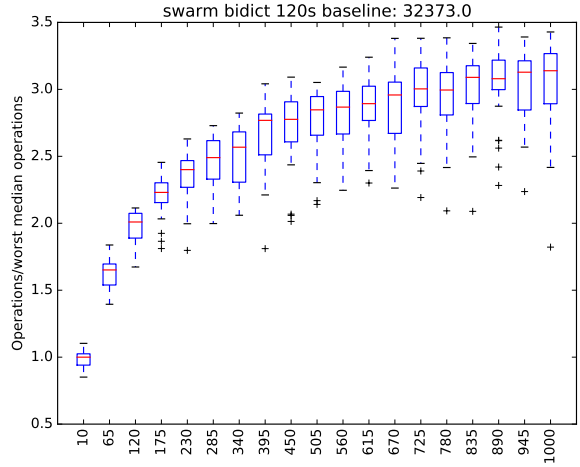

AVL 90s best: 218.0 @ 780
Branch coverage
120
140
160
180
200
220
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

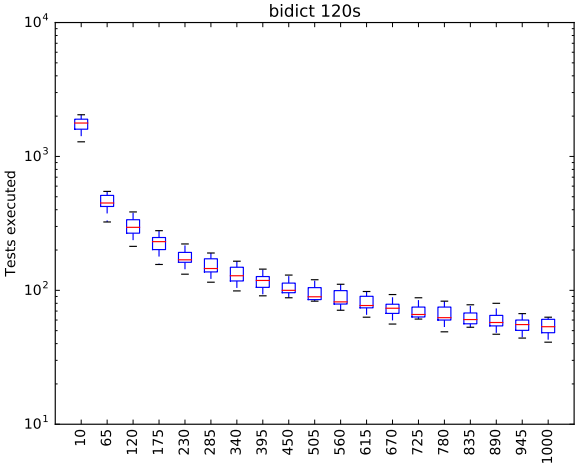

swarm AVL 90s best: 200.0 @ 120
Branch coverage
0
50
100
150
200
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

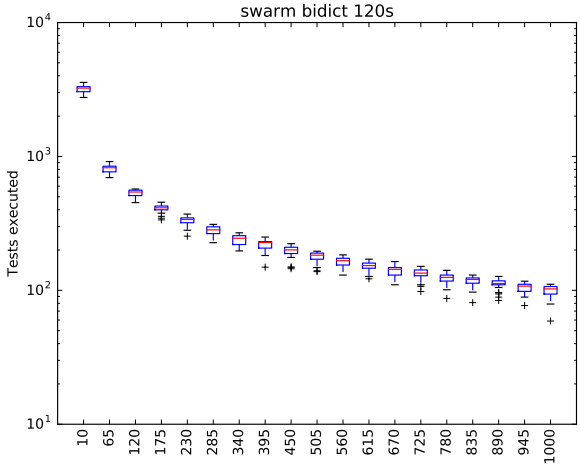

bidict 120s best: 373.0 @ 1000
Branch coverage
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

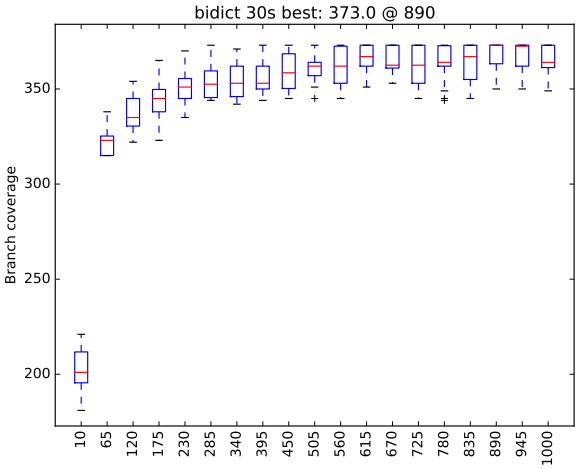

swarm bidict 120s best: 373.0 @ 1000
Branch coverage
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

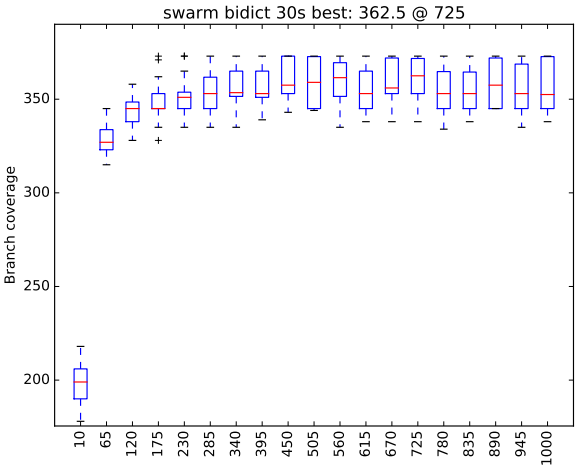

bidict 120s baseline: 17756.5
Operations/worst median operations
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

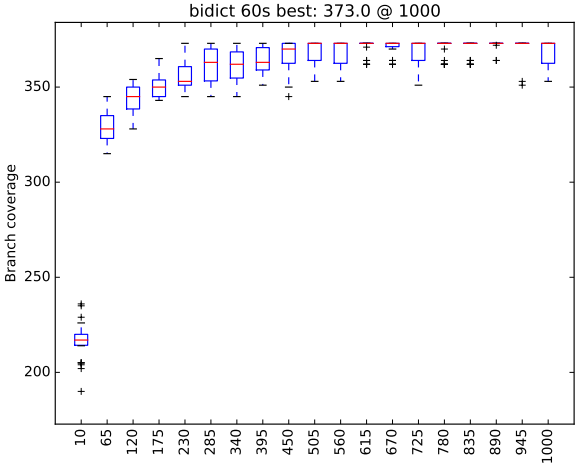

swarm bidict 120s baseline: 32373.0
Operations/worst median operations
0.5
1.0
1.5
2.0
2.5
3.0
3.5
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

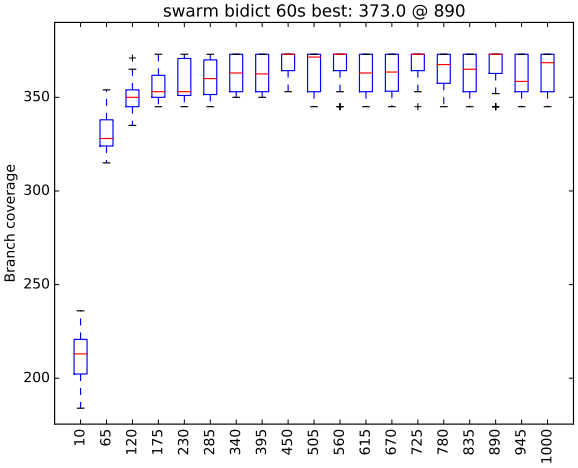

bidict 120s
Tests executed
10^1
10^2
10^3
10^4
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

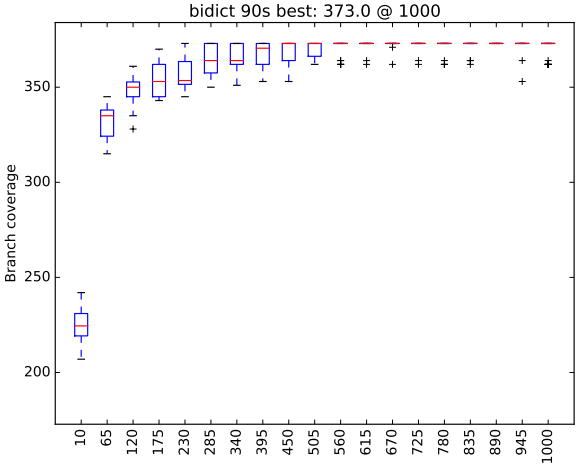

swarm bidict 120s
Tests executed
10^1
10^2
10^3
10^4
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

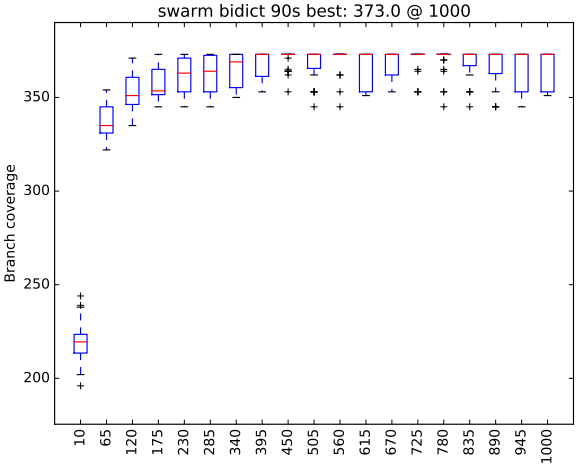

bidict 30s best: 373.0 @ 890
Branch coverage
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

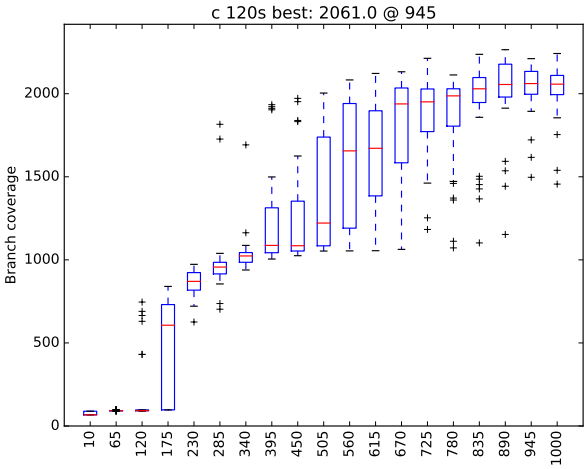

swarm bidict 30s best: 362.5 @ 725
Branch coverage
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

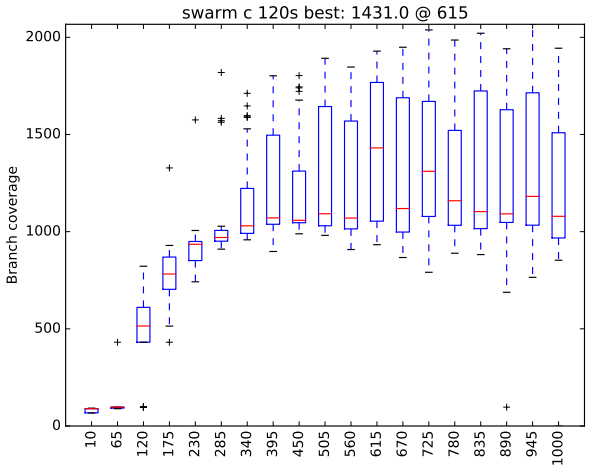

bidict 60s best: 373.0 @ 1000
Branch coverage
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

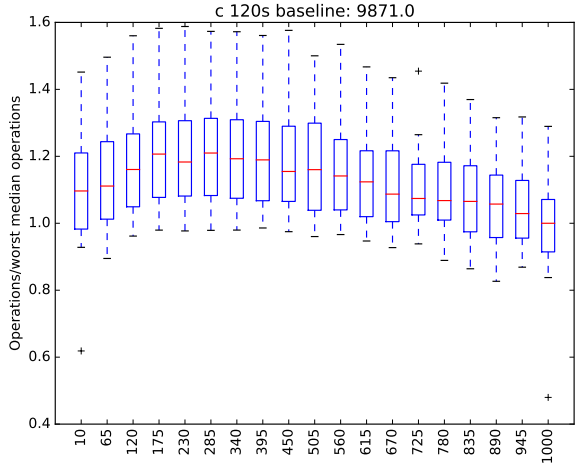

swarm bidict 60s best: 373.0 @ 890
Branch coverage
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

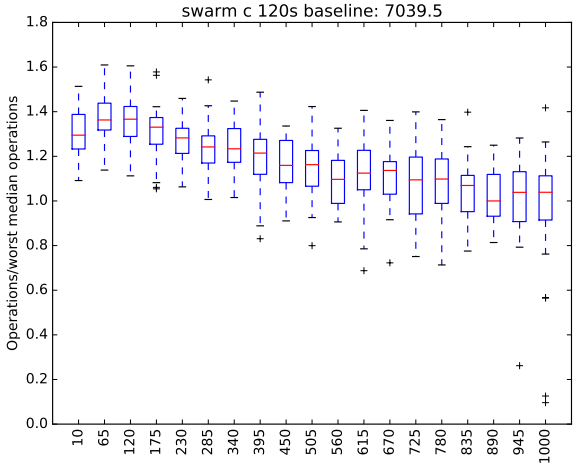

bidict 90s best: 373.0 @ 1000
Branch coverage
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

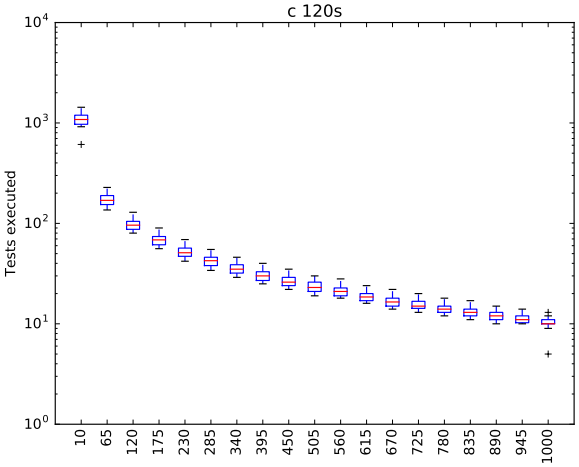

swarm bidict 90s best: 373.0 @ 1000
Branch coverage
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

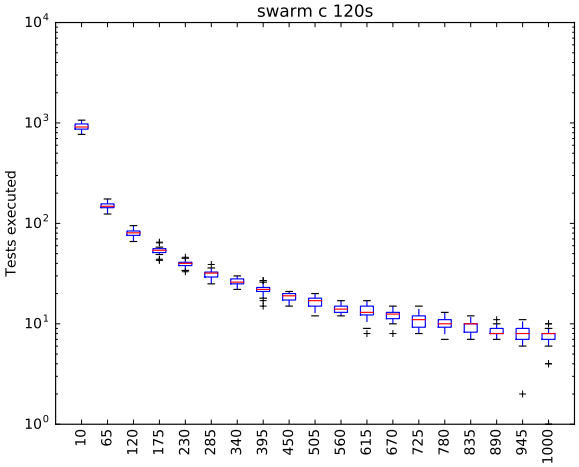

c 120s best: 2061.0 @ 945
Branch coverage
0
500
1000
1500
2000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

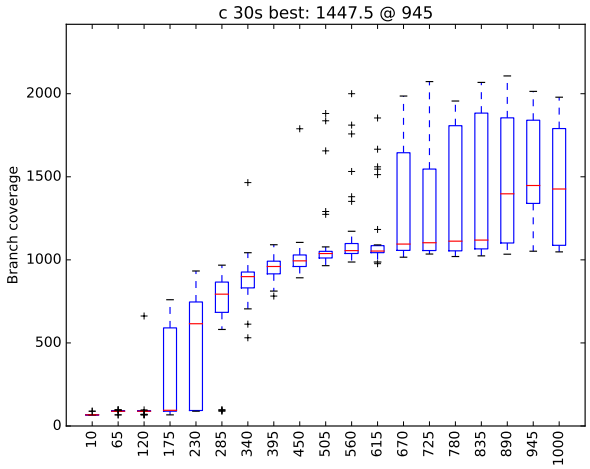

swarm c 120s best: 1431.0 @ 615
Branch coverage
0
500
1000
1500
2000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

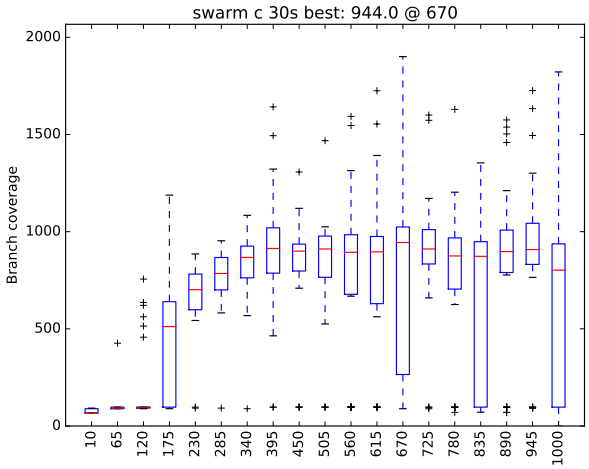

c 120s baseline: 9871.0
Operations/worst median operations
0.4
0.6
0.8
1.0
1.2
1.4
1.6
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

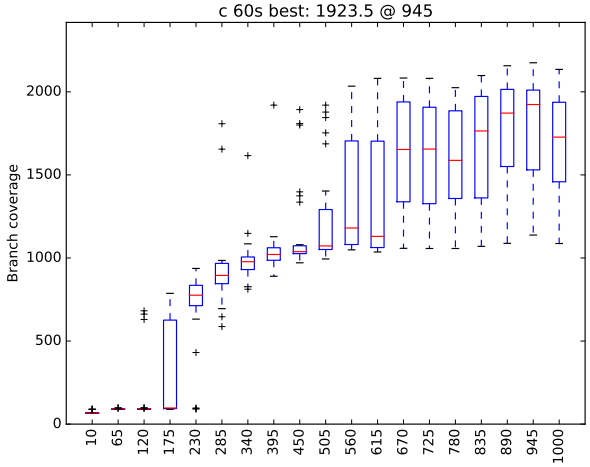

swarm c 120s baseline: 7039.5
Operations/worst median operations
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

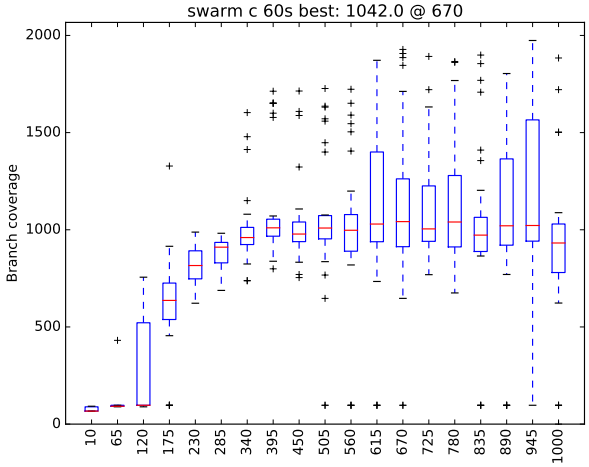

c 120s
Tests executed
10^0
10^1
10^2
10^3
10^4
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

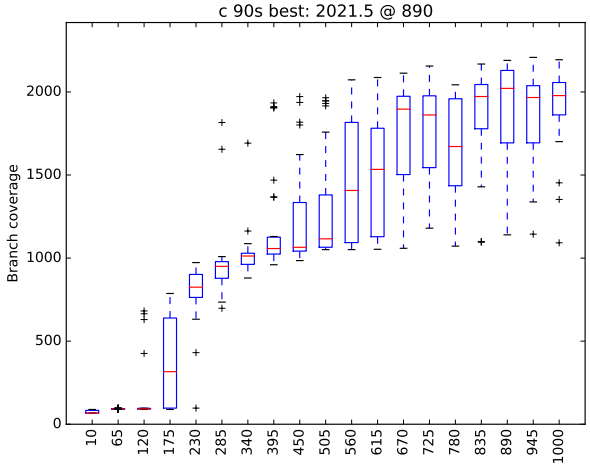

swarm c 120s
Tests executed
10^0
10^1
10^2
10^3
10^4
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

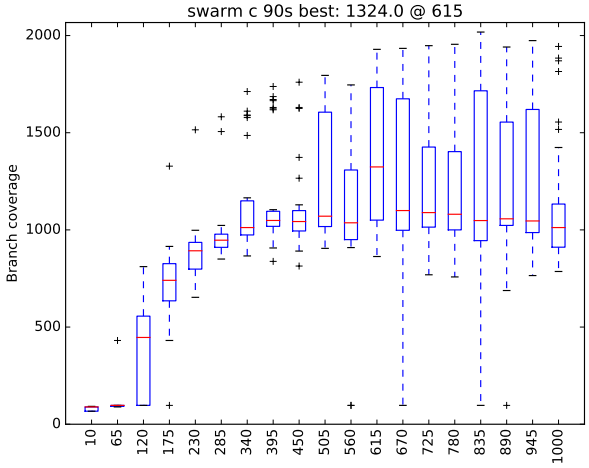

c 30s best: 1447.5 @ 945
Branch coverage
0
500
1000
1500
2000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

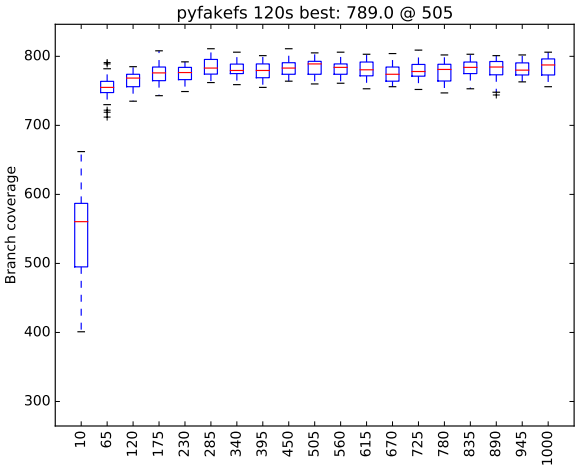

swarm c 30s best: 944.0 @ 670
Branch coverage
0
500
1000
1500
2000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

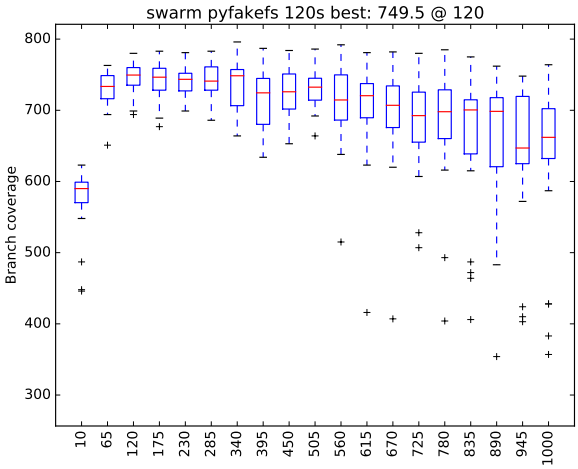

c 60s best: 1923.5 @ 945
Branch coverage
0
500
1000
1500
2000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

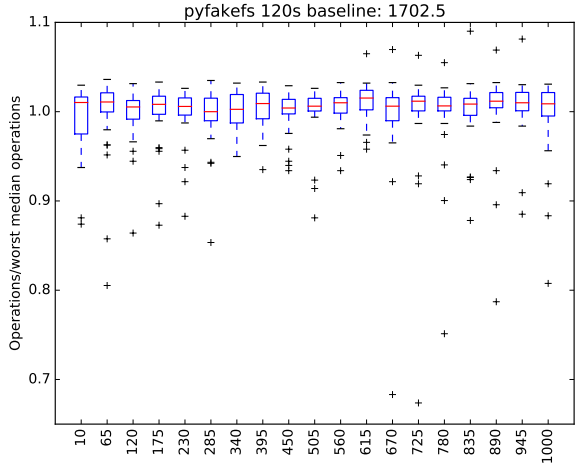

swarm c 60s best: 1042.0 @ 670
Branch coverage
0
500
1000
1500
2000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

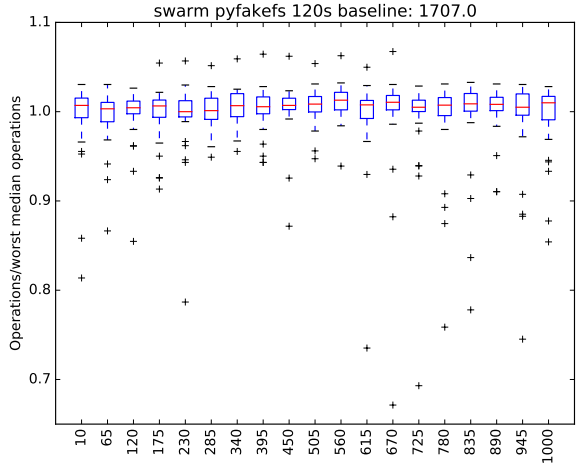

c 90s best: 2021.5 @ 890
Branch coverage
0
500
1000
1500
2000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

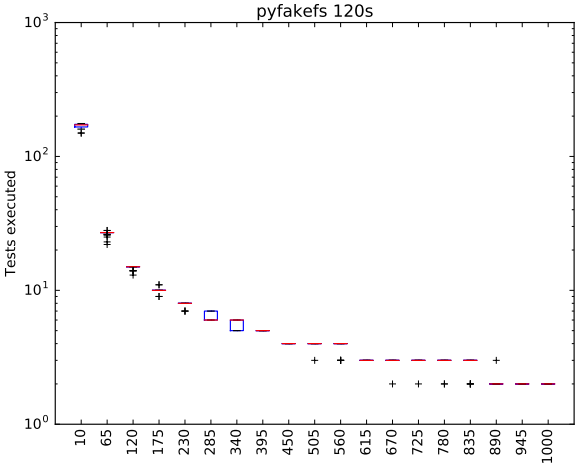

swarm c 90s best: 1324.0 @ 615
Branch coverage
0
500
1000
1500
2000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

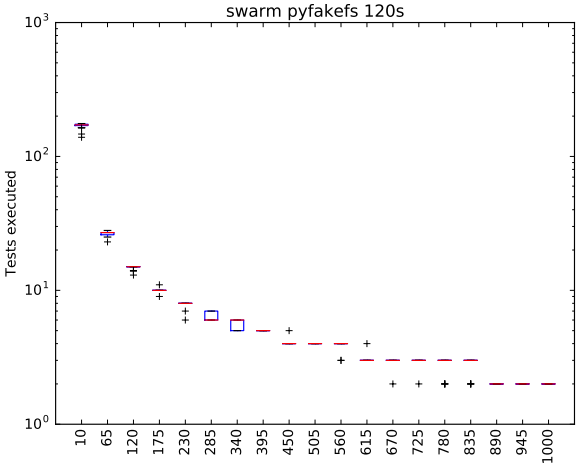

pyfakefs 120s best: 789.0 @ 505
Branch coverage
800
700
600
500
400
300
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

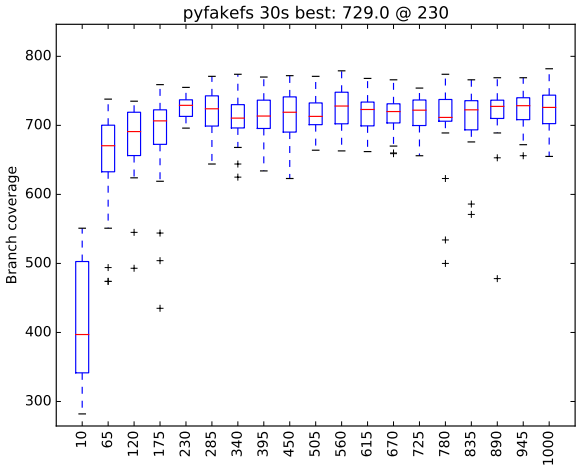

swarm pyfakefs 120s best: 749.5 @ 120
Branch coverage
800
700
600
500
400
300
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

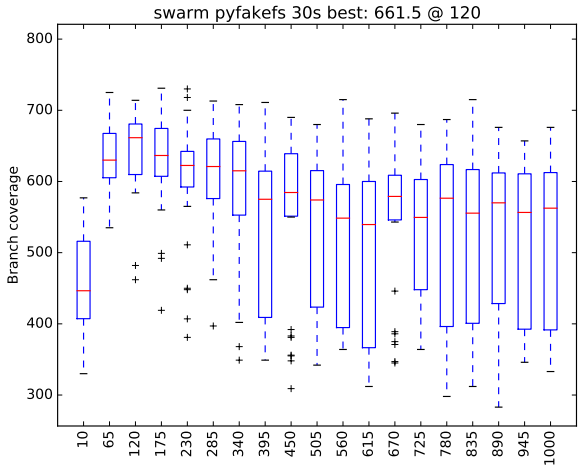

pyfakefs 120s baseline: 1702.5
Operations/worst median operations
1.1
1.0
0.9
0.8
0.7
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

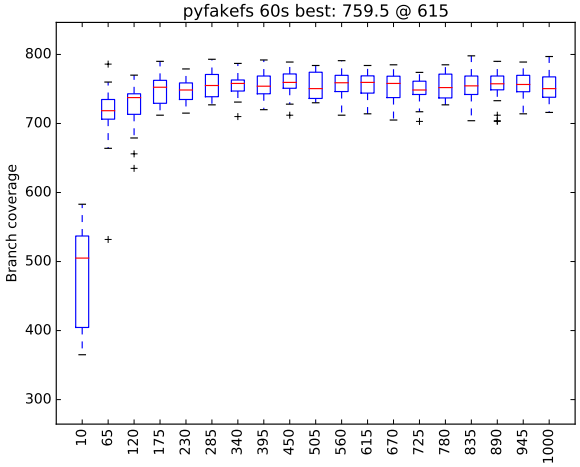

swarm pyfakefs 120s baseline: 1707.0
Operations/worst median operations
1.1
1.0
0.9
0.8
0.7
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

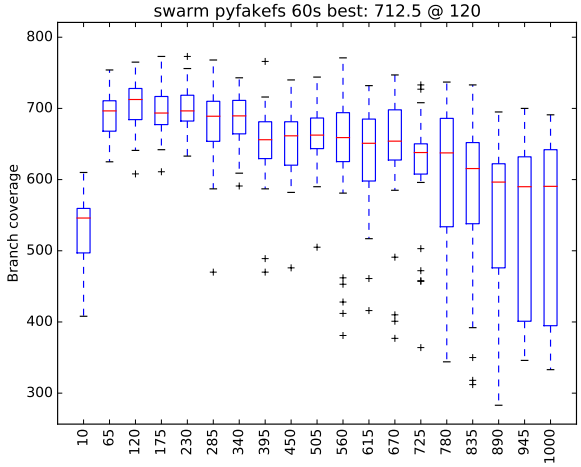

pyfakefs 120s
Tests executed
10^3
10^2
10^1
10^0
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

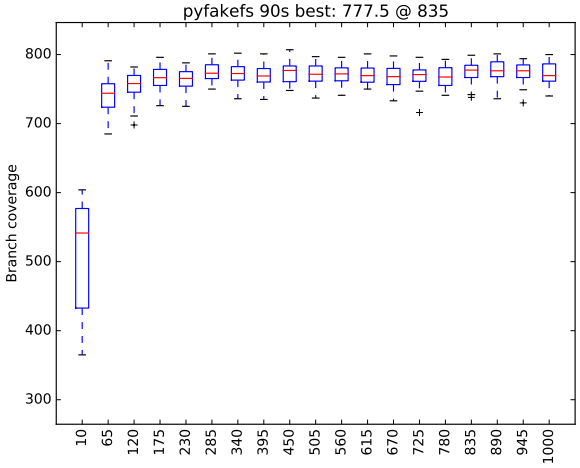

swarm pyfakefs 120s
Tests executed
10^3
10^2
10^1
10^0
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

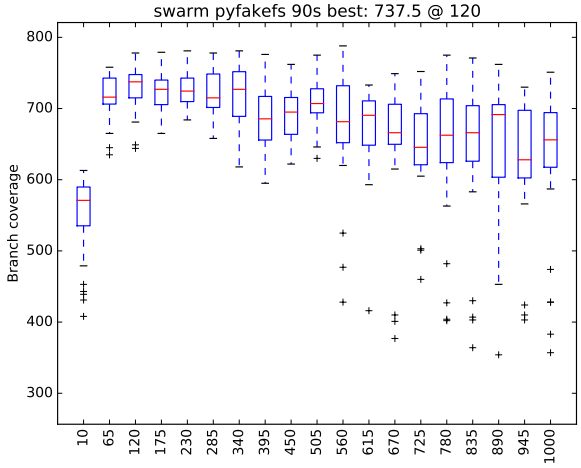

pyfakefs 30s best: 729.0 @ 230
Branch coverage
300
400
500
600
700
800
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

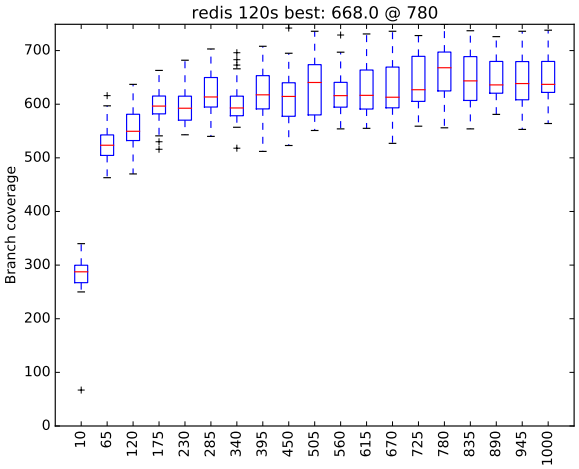

swarm pyfakefs 30s best: 661.5 @ 120
Branch coverage
300
400
500
600
700
800
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

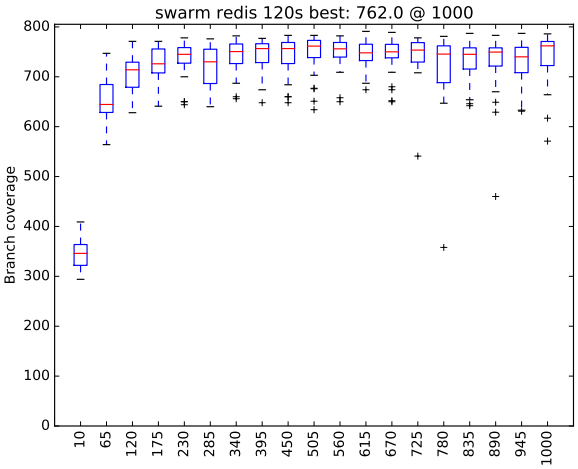

pyfakefs 60s best: 759.5 @ 615
Branch coverage
300
400
500
600
700
800
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

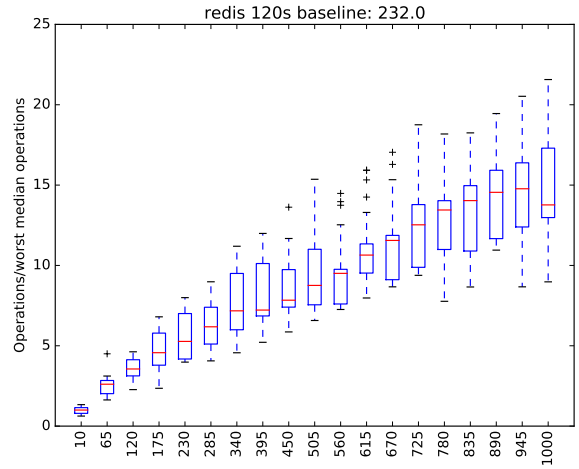

swarm pyfakefs 60s best: 712.5 @ 120
Branch coverage
300
400
500
600
700
800
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

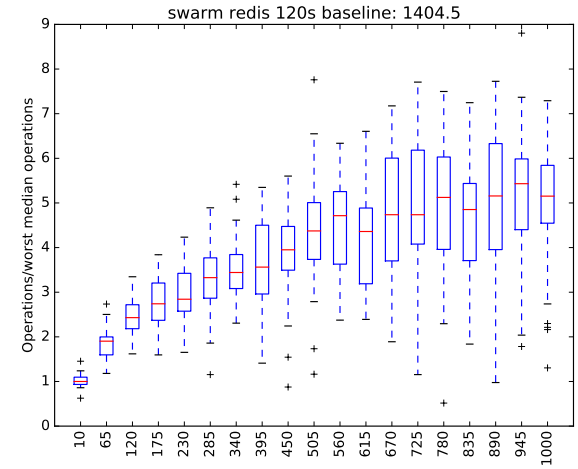

pyfakefs 90s best: 777.5 @ 835
Branch coverage
300
400
500
600
700
800
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

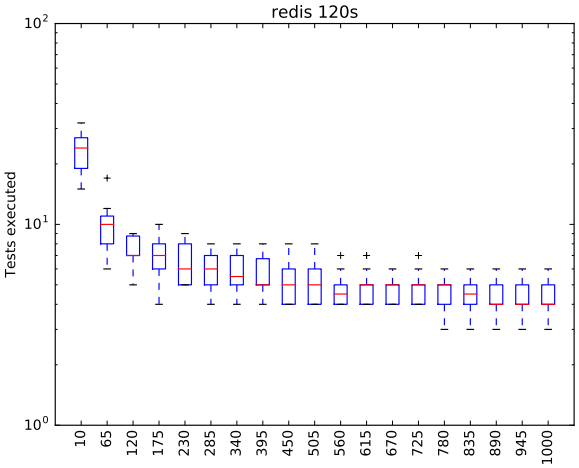

swarm pyfakefs 90s best: 737.5 @ 120
Branch coverage
300
400
500
600
700
800
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

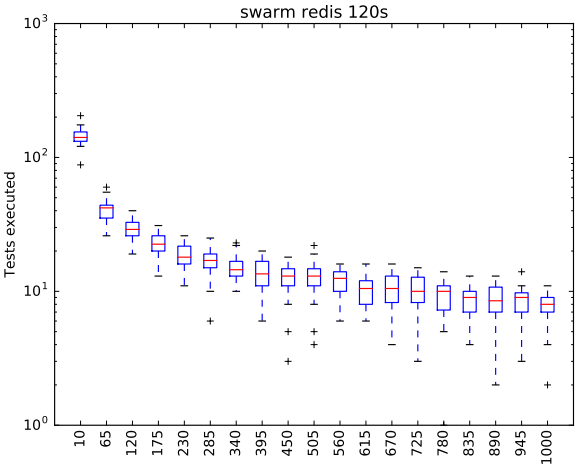

redis 120s best: 668.0 @ 780
Branch coverage
0
100
200
300
400
500
600
700
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

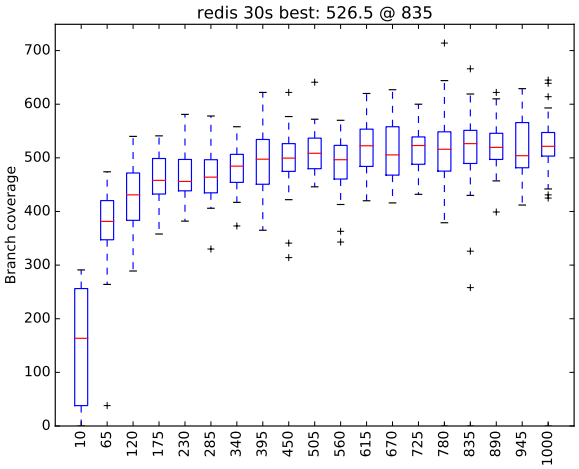

swarm redis 120s best: 762.0 @ 1000
Branch coverage
0
100
200
300
400
500
600
700
800
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

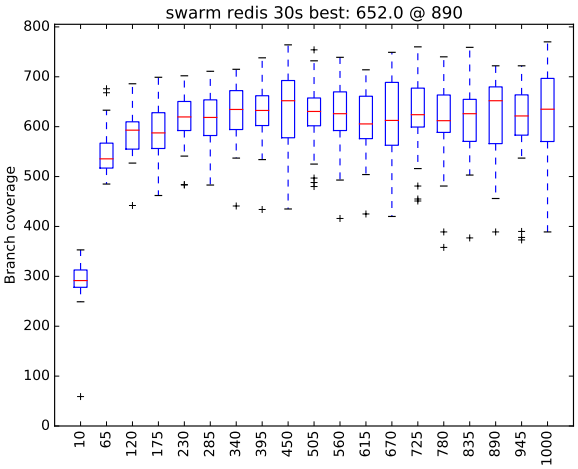

redis 120s baseline: 232.0
Operations/worst median operations
0
5
10
15
20
25
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

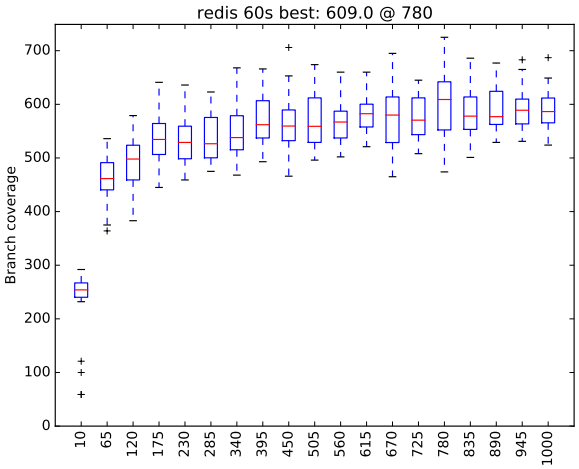

swarm redis 120s baseline: 1404.5
Operations/worst median operations
0
1
2
3
4
5
6
7
8
9
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

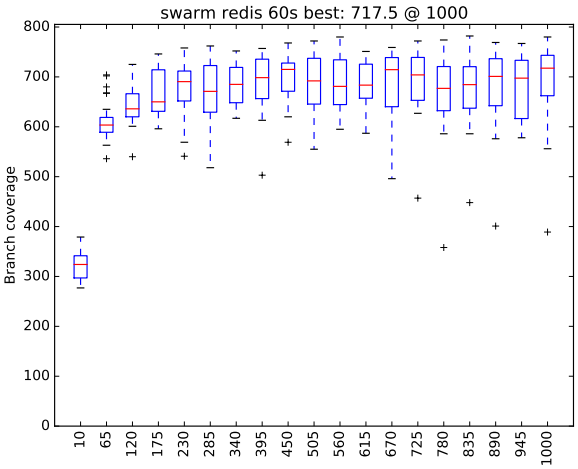

redis 120s
Tests executed
10^0
10^1
10^2
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

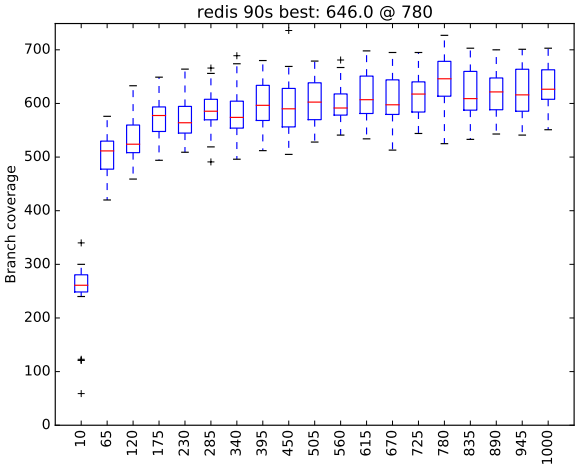

swarm redis 120s
Tests executed
10^0
10^1
10^2
10^3
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

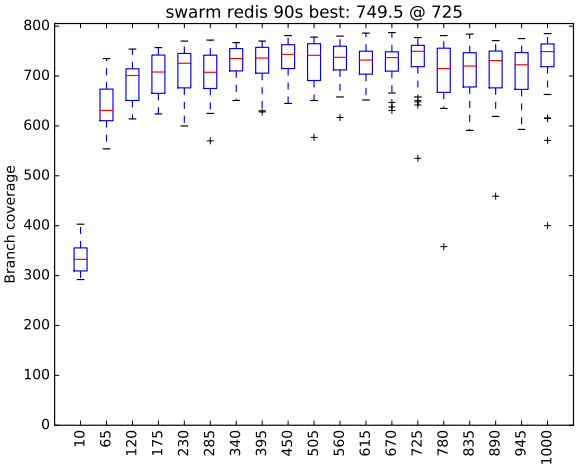

redis 30s best: 526.5 @ 835
Branch coverage
0
100
200
300
400
500
600
700
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

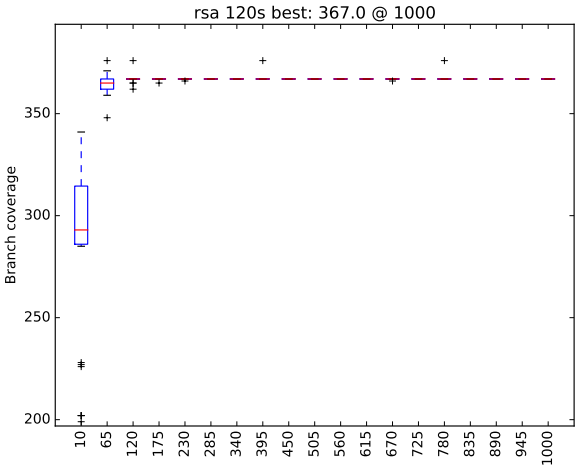

swarm redis 30s best: 652.0 @ 890
Branch coverage
0
100
200
300
400
500
600
700
800
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

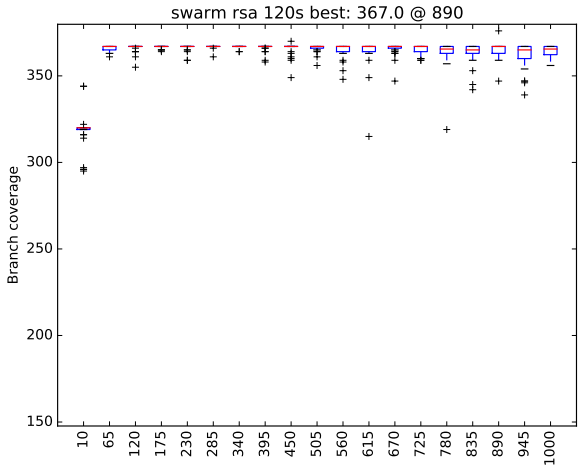

redis 60s best: 609.0 @ 780
Branch coverage
0
100
200
300
400
500
600
700
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

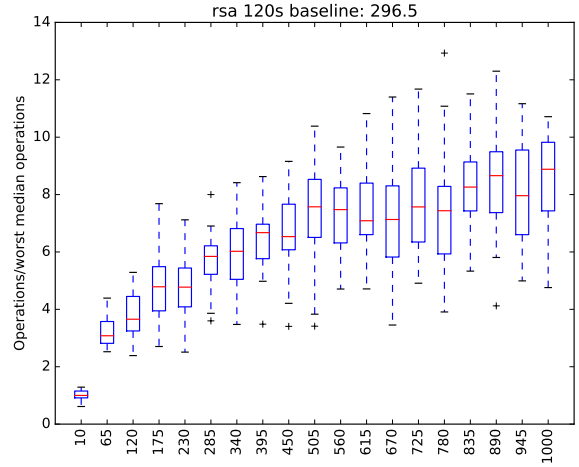

swarm redis 60s best: 717.5 @ 1000
Branch coverage
0
100
200
300
400
500
600
700
800
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

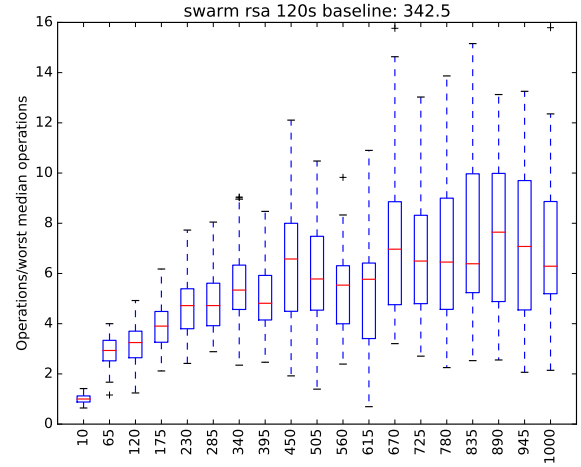

redis 90s best: 646.0 @ 780
Branch coverage
0
100
200
300
400
500
600
700
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

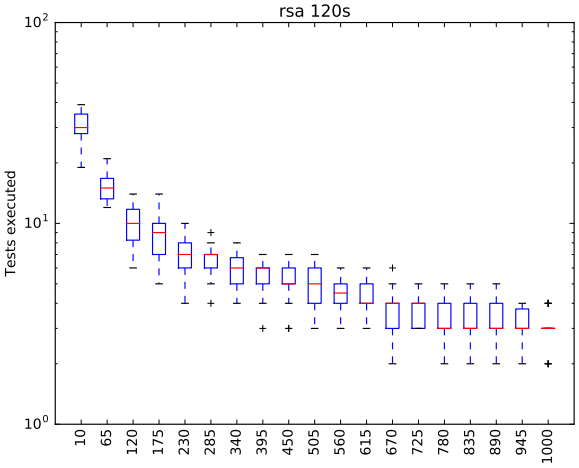

swarm redis 90s best: 749.5 @ 725
Branch coverage
0
100
200
300
400
500
600
700
800
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

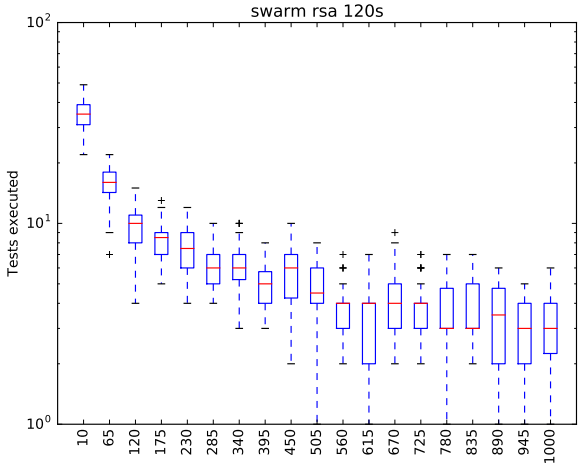

rsa 120s best: 367.0 @ 1000
Branch coverage
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

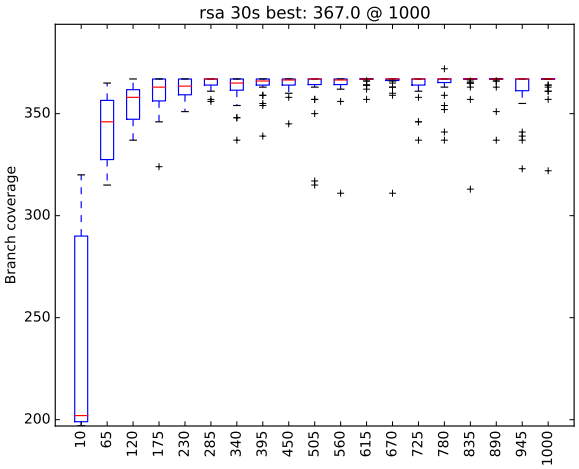

swarm rsa 120s best: 367.0 @ 890
Branch coverage
150
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

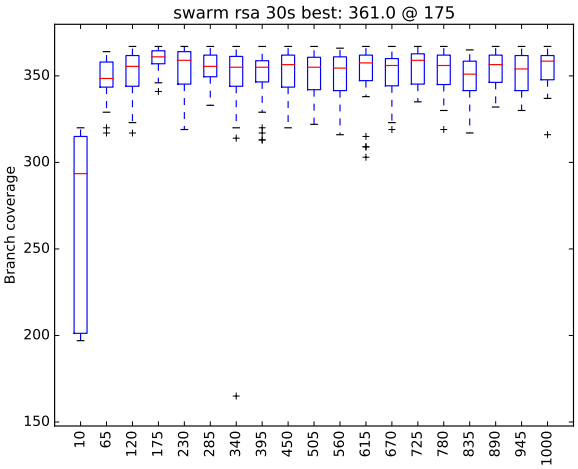

rsa 120s baseline: 296.5
Operations/worst median operations
0 2 4 6 8 10 12 14
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

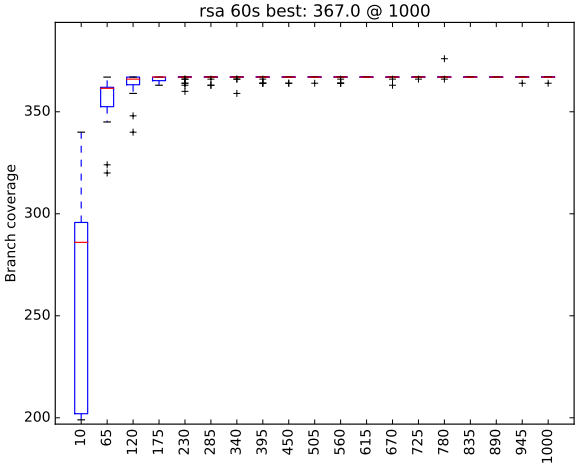

swarm rsa 120s baseline: 342.5
Operations/worst median operations
0 2 4 6 8 10 12 14 16
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

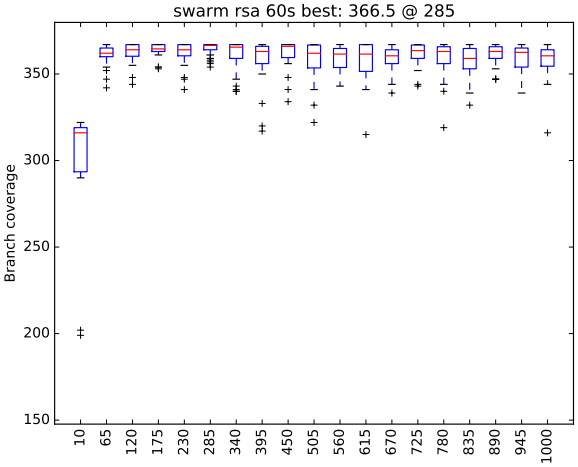

rsa 120s
Tests executed
10^0 10^1 10^2
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

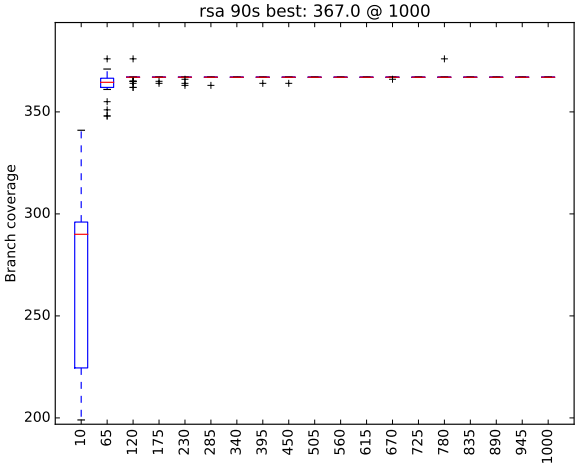

swarm rsa 120s
Tests executed
10^0 10^1 10^2
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

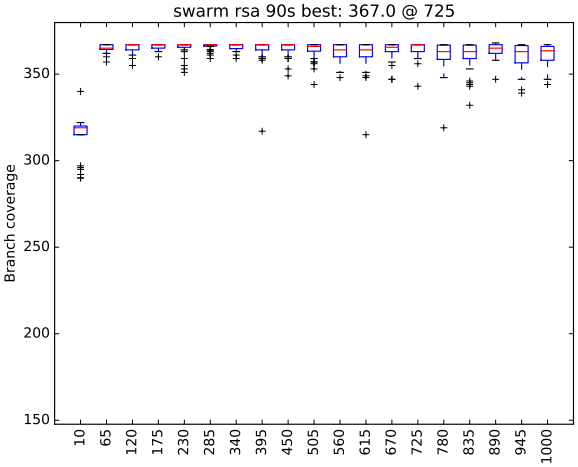

rsa 30s best: 367.0 @ 1000
Branch coverage
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

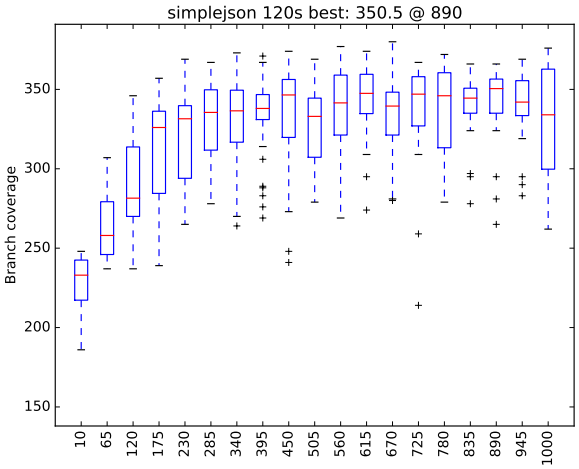

swarm rsa 30s best: 361.0 @ 175
Branch coverage
150
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

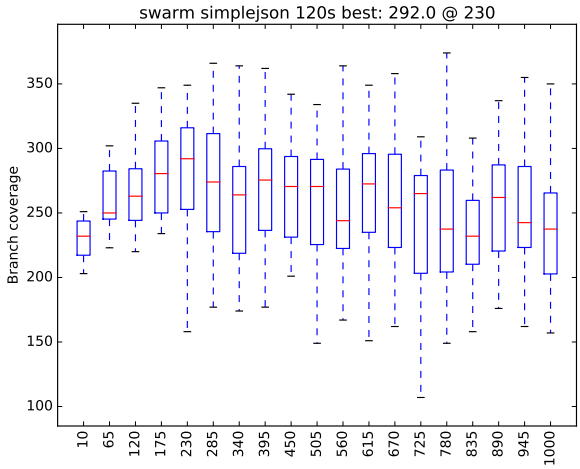

rsa 60s best: 367.0 @ 1000
Branch coverage
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

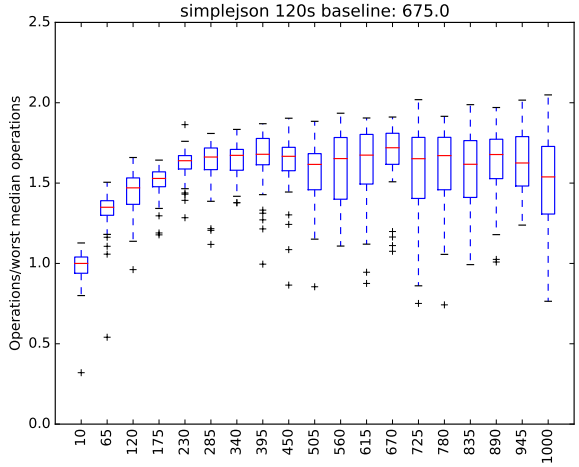

swarm rsa 60s best: 366.5 @ 285
Branch coverage
150
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

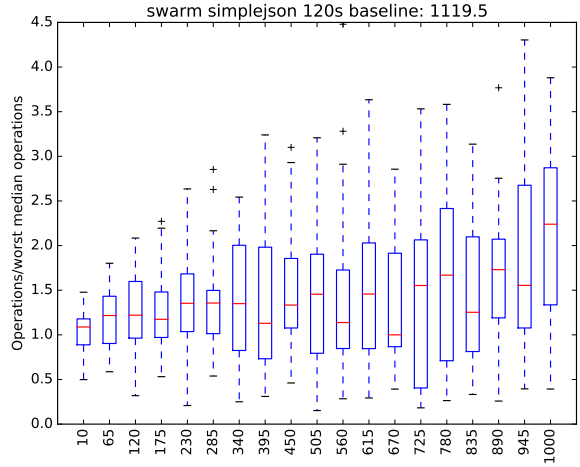

rsa 90s best: 367.0 @ 1000
Branch coverage
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

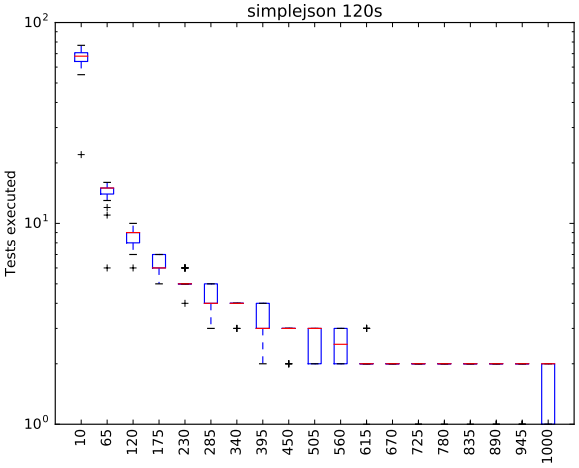

swarm rsa 90s best: 367.0 @ 725
Branch coverage
150
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

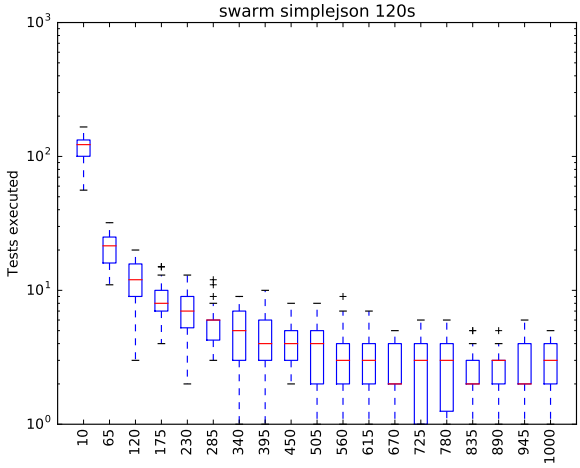

simplejson 120s best: 350.5 @ 890
Branch coverage
150
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

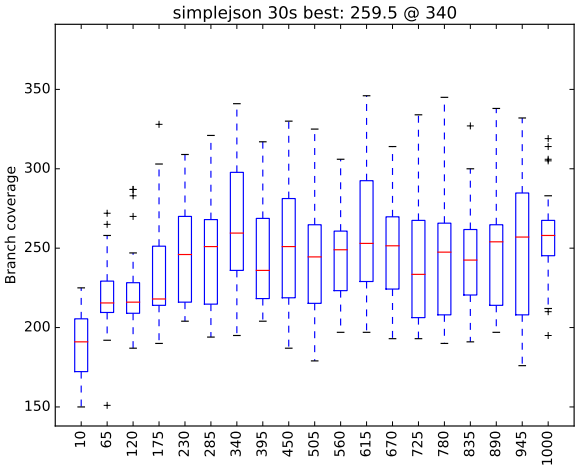

swarm simplejson 120s best: 292.0 @ 230
Branch coverage
100
150
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

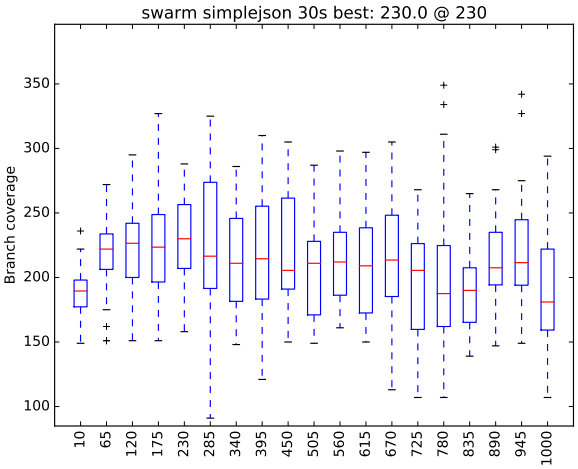

simplejson 120s baseline: 675.0
Operations/worst median operations
0.0
0.5
1.0
1.5
2.0
2.5
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

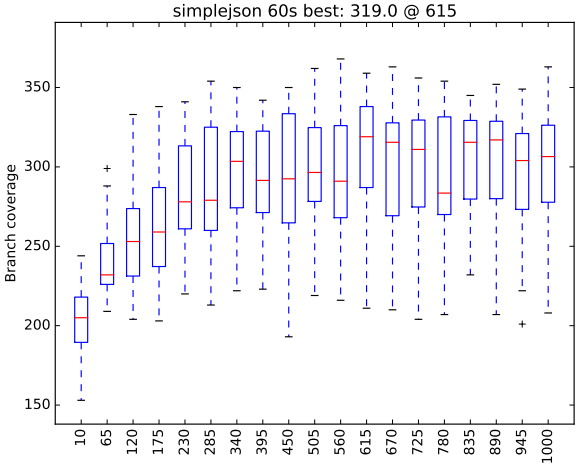

swarm simplejson 120s baseline: 1119.5
Operations/worst median operations
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

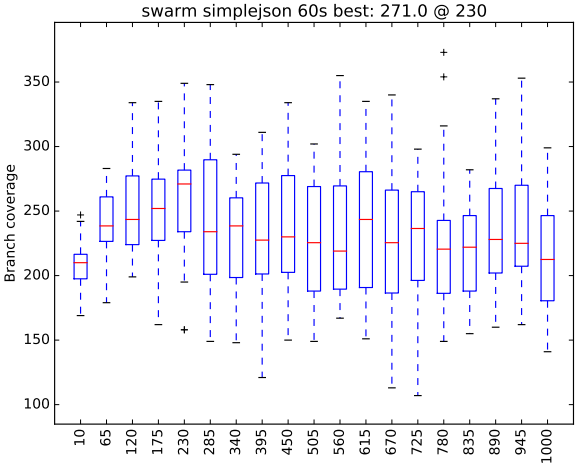

simplejson 120s
Tests executed
10^0
10^1
10^2
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

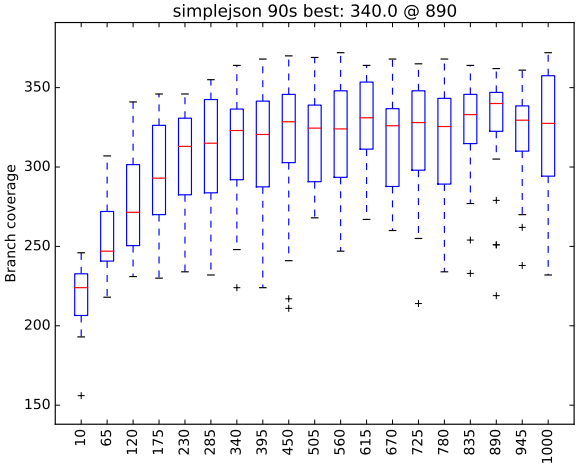

swarm simplejson 120s
Tests executed
10^0
10^1
10^2
10^3
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

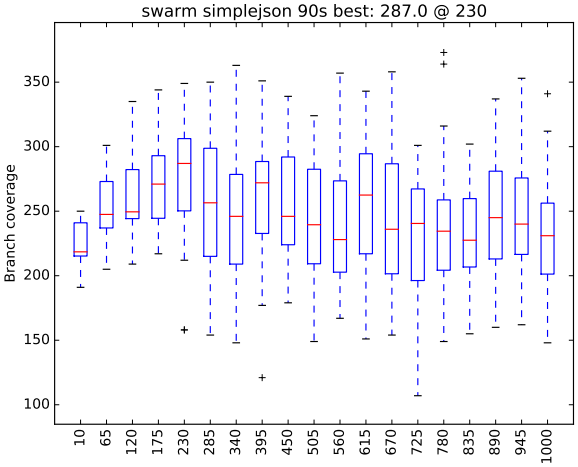

simplejson 30s best: 259.5 @ 340
Branch coverage
150
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

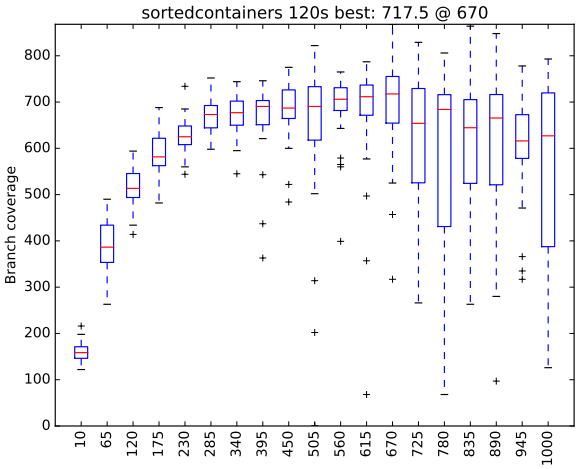

swarm simplejson 30s best: 230.0 @ 230
Branch coverage
100
150
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

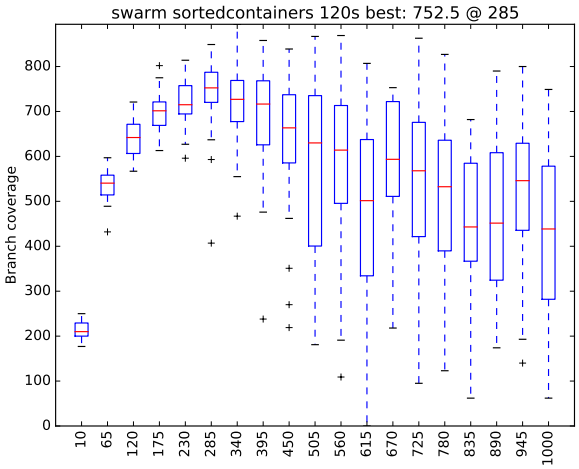

simplejson 60s best: 319.0 @ 615
Branch coverage
150
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

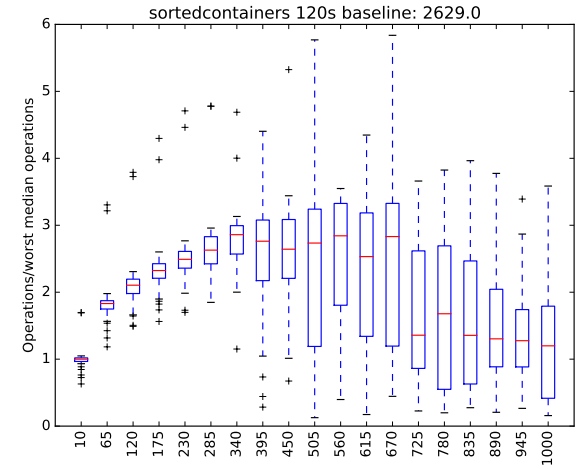

swarm simplejson 60s best: 271.0 @ 230
Branch coverage
100
150
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

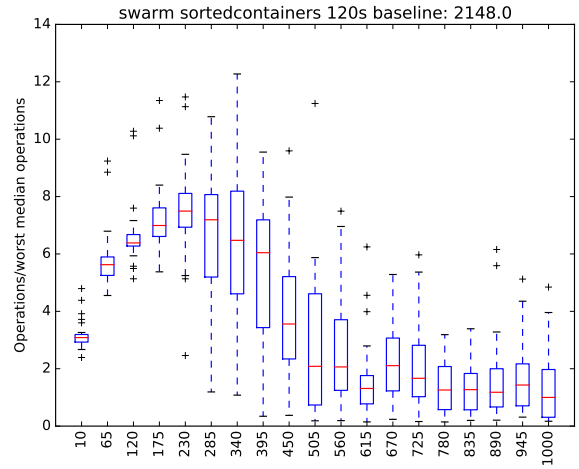

simplejson 90s best: 340.0 @ 890
Branch coverage
150
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

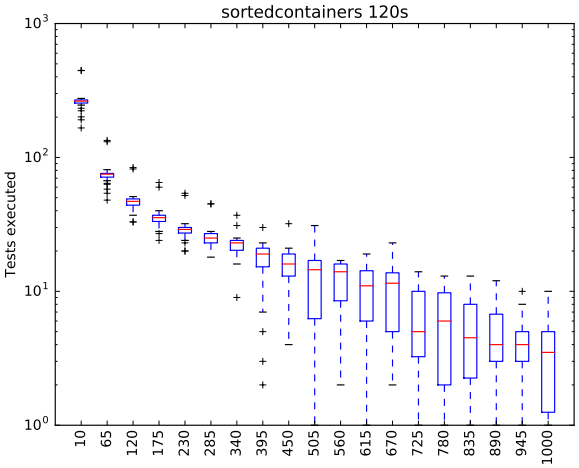

swarm simplejson 90s best: 287.0 @ 230
Branch coverage
100
150
200
250
300
350
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

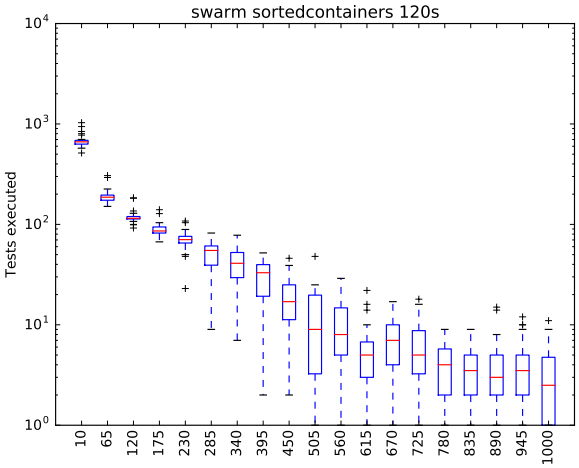

sortedcontainers 120s best: 717.5 @ 670
Branch coverage
0
100
200
300
400
500
600
700
800
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

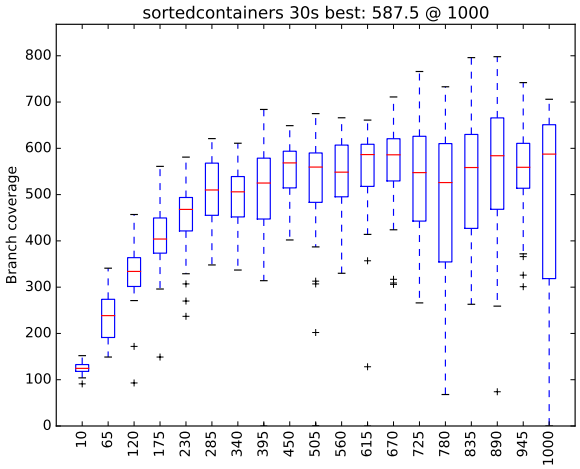

swarm sortedcontainers 120s best: 752.5 @ 285
Branch coverage
0
100
200
300
400
500
600
700
800
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

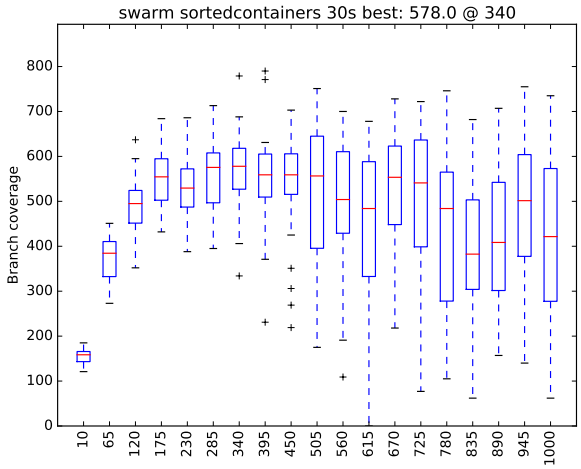

sortedcontainers 120s baseline: 2629.0
Operations/worst median operations
0
1
2
3
4
5
6
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

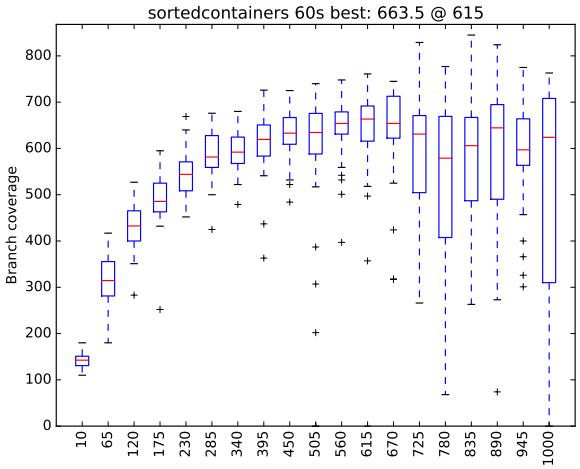

swarm sortedcontainers 120s baseline: 2148.0
Operations/worst median operations
0
2
4
6
8
10
12
14
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

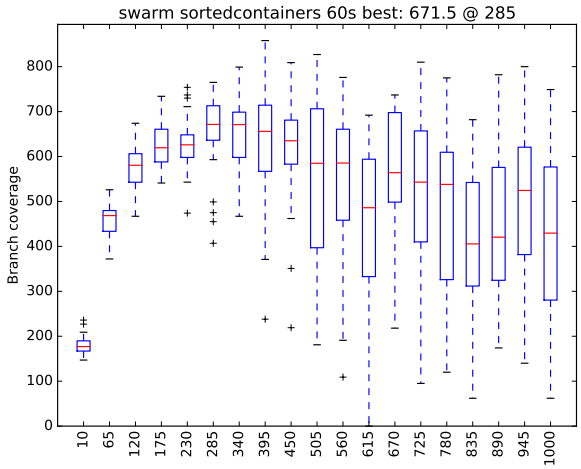

sortedcontainers 120s
Tests executed
10^0
10^1
10^2
10^3
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

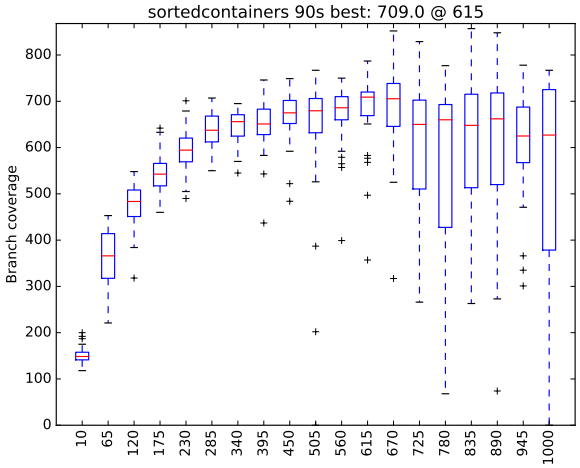

swarm sortedcontainers 120s
Tests executed
10^0
10^1
10^2
10^3
10^4
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

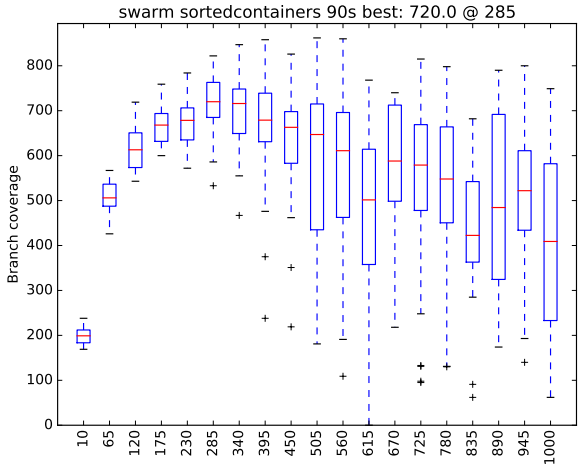

sortedcontainers 30s best: 587.5 @ 1000
Branch coverage
0
100
200
300
400
500
600
700
800
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

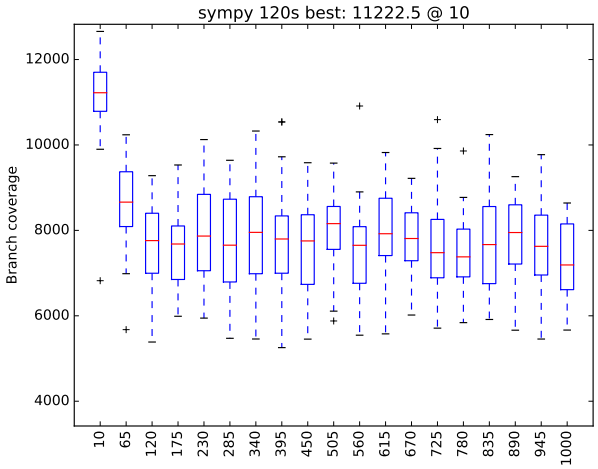

swarm sortedcontainers 30s best: 578.0 @ 340
Branch coverage
0
100
200
300
400
500
600
700
800
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

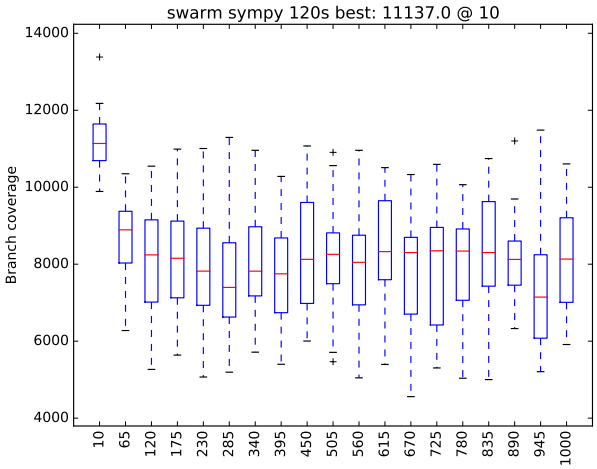

sortedcontainers 60s best: 663.5 @ 615
Branch coverage
0
100
200
300
400
500
600
700
800
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

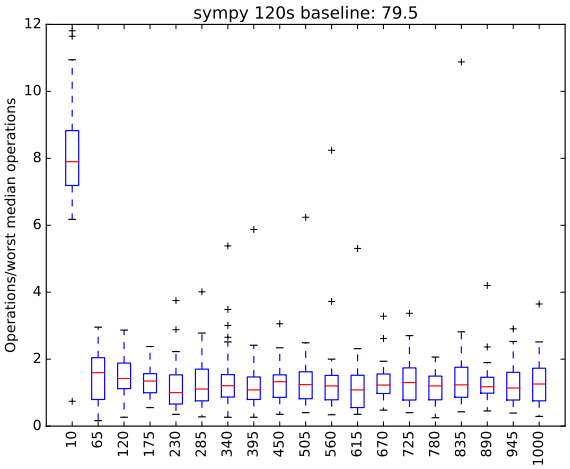

swarm sortedcontainers 60s best: 671.5 @ 285
Branch coverage
0
100
200
300
400
500
600
700
800
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

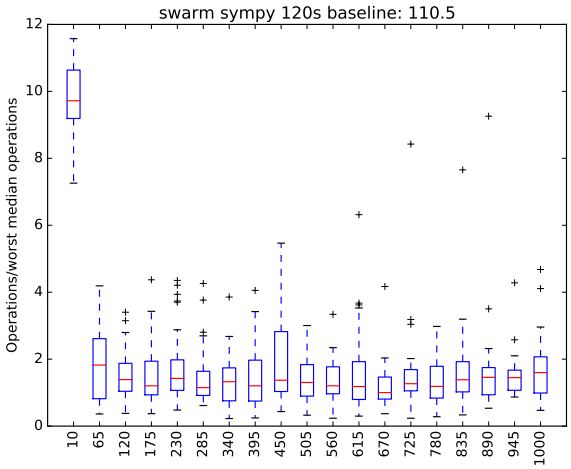

sortedcontainers 90s best: 709.0 @ 615
Branch coverage
0
100
200
300
400
500
600
700
800
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

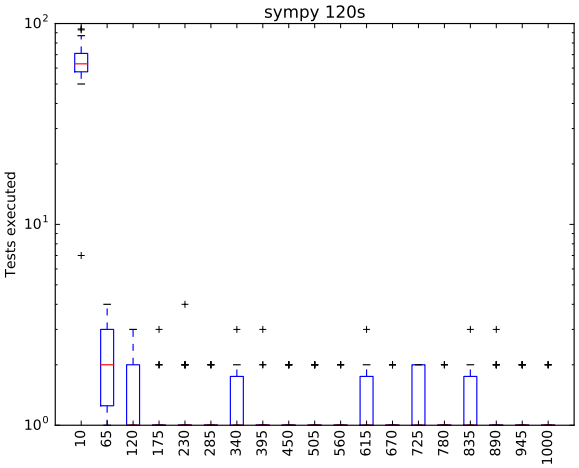

swarm sortedcontainers 90s best: 720.0 @ 285
Branch coverage
0
100
200
300
400
500
600
700
800
10
65
120
175
230
285
340
395
450
505
560
615
670
725
780
835
890
945
1000

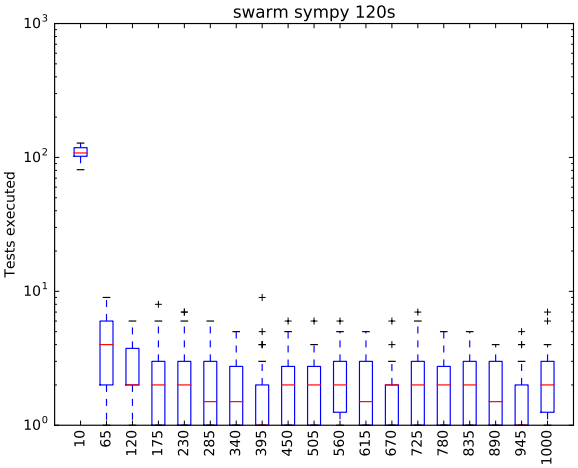

sympy 120s best: 11222.5 @ 10
Branch coverage
12000
10000
8000
6000
4000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

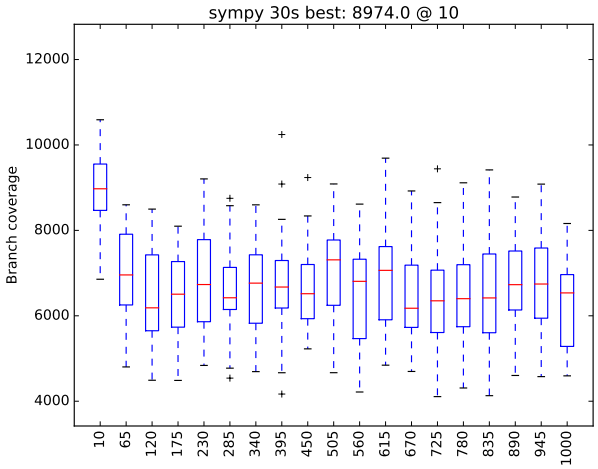

swarm sympy 120s best: 11137.0 @ 10
Branch coverage
14000
12000
10000
8000
6000
4000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

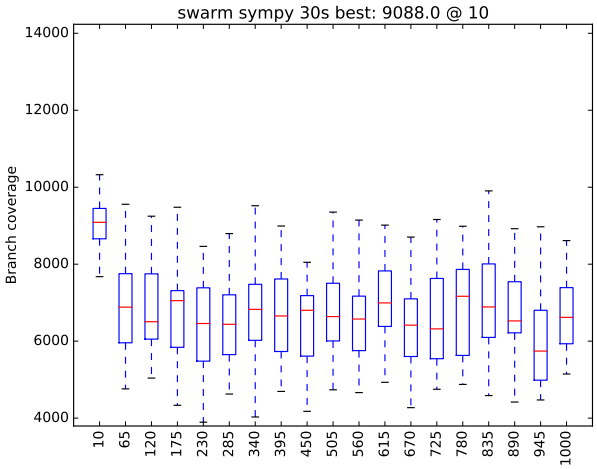

sympy 120s baseline: 79.5
Operations/worst median operations
12
10
8
6
4
2
0
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

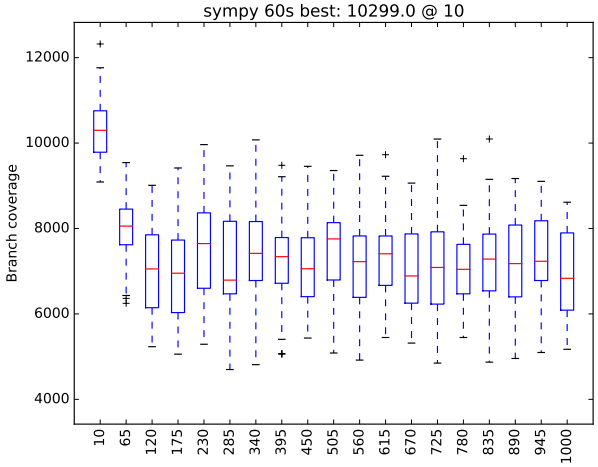

swarm sympy 120s baseline: 110.5
Operations/worst median operations
12
10
8
6
4
2
0
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

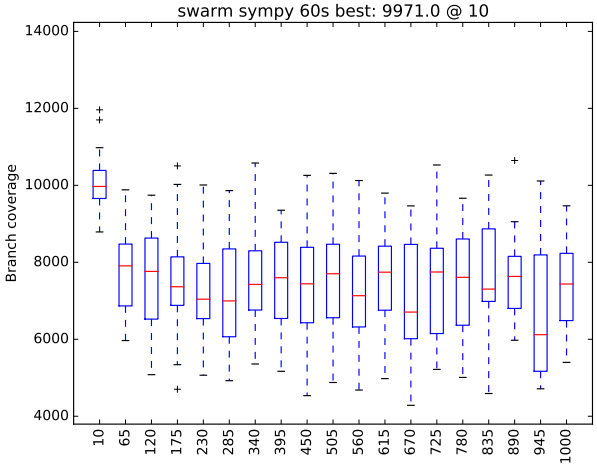

sympy 120s
Tests executed
10^2
10^1
10^0
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

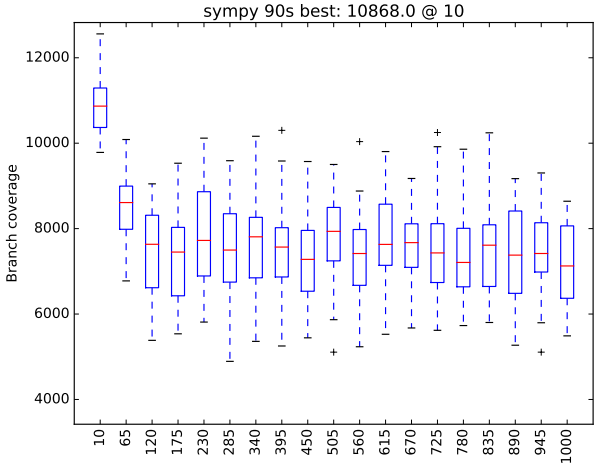

swarm sympy 120s
Tests executed
10^3
10^2
10^1
10^0
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000

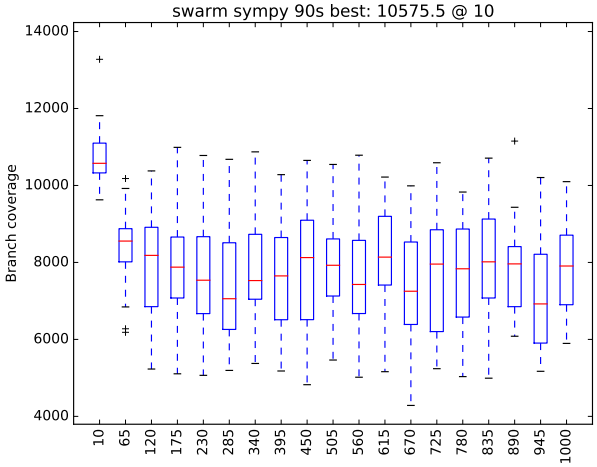

sympy 30s best: 8974.0 @ 10
Branch coverage
4000
6000
8000
10000
12000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000


swarm sympy 30s best: 9088.0 @ 10
Branch coverage
4000
6000
8000
10000
12000
14000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000


sympy 60s best: 10299.0 @ 10
Branch coverage
4000
6000
8000
10000
12000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000


swarm sympy 60s best: 9971.0 @ 10
Branch coverage
4000
6000
8000
10000
12000
14000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000


sympy 90s best: 10868.0 @ 10
Branch coverage
4000
6000
8000
10000
12000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000


swarm sympy 90s best: 10575.5 @ 10
Branch coverage
4000
6000
8000
10000
12000
14000
10 65 120 175 230 285 340 395 450 505 560 615 670 725 780 835 890 945 1000