PB16
Pseudo-Boolean Competition 2016

Olivier ROUSSEL

19th International Conference
on Theory and Applications of Satisfiability Testing,
SAT'16

June 5, 2016

What's a competition worth?

The goal of a competition is to:

- evaluate solvers in the same conditions
- help collecting publicly available benchmarks
- help identifying new solvers on the market
- help the community identify good ideas and strange results: the goal is to raise questions and get new ideas!

Competitions should not be misunderstood:

- The results are not an absolute truth: they depend on the benchmark selection, experimental condition,...
- A competition is not limited to a ranking: rankings are just an over-simplified view, but still relevant to motivate authors
- There are a lot of data collected and published to benefit the whole community
- Competitions must be driven by the community: benchmark submission/selection advices, suggestions for improvements...

Pseudo-Boolean Constraints

- **Linear** (LIN) pseudo-Boolean (PB) constraint = sum of integer $\times$ a literal $\{\geq, \leq, =\}$ constant
  Example: $3x_1 - 3x_2 + 2\bar{x}_3 + \bar{x}_4 + x_5 \geq 5$
- **Non-linear** (NLC) pseudo-Boolean (PB) constraint = sum of integer $\times$ a product of literals $\{\geq, \leq, =\}$ constant
  Example: $3x_1\bar{x}_2 - 3x_2x_4 + 2\bar{x}_3 + \bar{x}_4 + x_5x_6x_7 \geq 5$
- As an example, PB allows compact encodings of:
  - cardinalities: $x_1 + x_2 + x_3 \geq 2$
  - adder (C=A+B): $2c_1 + c_0 = 2a_1 + a_0 + 2b_1 + b_0$
  - knapsack:
    $max : 5x_1 + 10x_2 + 2x_3;$
    $5x_1 + 8x_2 + x_3 \leq 10$
- Cutting-planes proof system stronger than resolution: PHP easily solved in polynomial time

Different problems: PBS, PBO, WBO

- **PBS (Pseudo Boolean Satisfaction)**
  decide of the satisfiability of a conjunction of PB constraints (decision problem)
- **PBO (Pseudo Boolean Optimization)**
  find a model of a conjunction of PB constraints which optimizes one objective function
- **WBO (Weighted Boolean Optimization)**
  - maximum satisfiability for PB constraints
  - hard constraints **must** be satisfied
  - soft constraints may be violated, but this has a cost
  - as in WCSP, there is a top cost. Interpretations with a cost greater or equal to the top cost are non admissible.

Benchmark categories (1)

For PBS/PBO, classification based on the objective function

- DEC: No objective function to optimize (decision problem). The solver must simply find a solution.
- OPT: An objective function is present. The solver must find a solution with the best possible value of the objective function.

For WBO, classification based on the existence of hard clauses

- SOFT: No hard clause at all.
- PARTIAL: At least one hard clause.

Benchmark categories (2)

Classification based on the size of coefficients

SMALLINT small integers: no constraint with a sum of coefficients greater than 2^{20} (20 bits): expected to be safe for solvers using 32 bits integers and simple techniques (be careful with learning), but strong limit to the encoding of concrete problems.

BIGINT big integers: at least one constraint with a sum of coefficients greater than 2^{20} (20 bits): requires arbitrary precision.

Classification based on the linearity of constraints

LIN All constraints are linear

NLC At least one constraint is non linear (contains products of literals)

New instances (?)

PBS-PBO

- 1048 instances from PB15eval (submitted by Jan Elffers, or obtained on the web)
- 226 instances from LION9 challenge: Development Assurance Level problems (multi-objective translated to single objective)
- 180 instances from PB instance generator for SHA-1 by Vegard Nossum (https://github.com/vegard/sha1-sat), 21 to 23 rounds, 80 to 160 bits.
- 304 instances from “Finding Synchronization Codes to Boost Compression by Substring Enumeration”, Dany Vohl, Claude-Guy Quimper, and Danny Dubé, ModRef 2012

WBO

- nothing new!

Instance selection

- Basically no selection!
- The idea was to let the user generate his own selection and obtain his own competition results on that selection. Requires a new web site (not yet implemented).
- Add all instances from PB12, PB11, PB10 to allow comparison with solvers of these competitions (requires a new web site)
- Add most instances from the PB15 evaluation

Solvers

Submitted solvers:

- cdcl-cuttingplanes (Jan ELFFERS)
- NaPS (Masaihko SAKAI)
- Open-WBO (Ruben MARTINS)
- Sat4J (Daniel LE BERRE and Anne PARRAIN)
- toysat (Masahiro SAKAI)

Unsubmitted solvers:

- minisatp (Niklas EEN, Niklas SORENSSON)
  2012 version available on GitHub

Categories and selected instances

- **DEC-SMALLINT-LIN (1783 instances)**
- DEC-SMALLINT-NLC (100 instances)
- DEC-BIGINT-LIN (37 instances)
- DEC-BIGINT-NLC (0 instance)
- **OPT-SMALLINT-LIN (1600 instances)**
- **OPT-SMALLINT-NLC (506 instances)**
- **OPT-BIGINT-LIN (1109 instances)**
- OPT-BIGINT-NLC (40 instances)
- **PARTIAL-SMALLINT-LIN (549 instances)**
- **PARTIAL-BIGINT-LIN (263 instances)**
- **SOFT-SMALLINT-LIN (201 instances)**
- SOFT-BIGINT-LIN (46 instances)

Evaluation environment

kindly provided by CRIL, University of Artois, France

- Same environment as PB12
- Cluster of bi-Xeon quad-core 2.66 GHz, 8 MB cache, 32 GB RAM
- Each solver was given a time limit of 30 minutes (1800s) and a memory limit of 15500 MB (to avoid swapping).
- 2 solvers per node (each solver is given 4 cores)
  - limited interactions because of the 2 CPU and the memory limit
  - *only two instances of the same solver allowed to run concurrently*
- 364 days of CPU time used

Verification of results

- The environment performs the following, efficient checks:
  - for SATISFIABLE answers, solvers must output a complete instantiation and the system checks that it satisfies all constraints
  - for UNSATISFIABLE answers, the system only checks that no other solver proved satisfiability
  - for OPTIMUM FOUND answers, solvers must output a complete instantiation; the system checks if all constraints are satisfied and that no other solver found a better solution
- UNSATISFIABLE and OPTIMUM FOUND answers cannot be completely checked efficiently and therefore should be taken with caution.
- Solvers giving a wrong answer in a category are disqualified in that category.

Ranking of solvers and Virtual Best Solver (VBS)

Ranking based on two criteria:

1. the number of solved instances
2. ties are broken by considering the cumulated time on solved instances

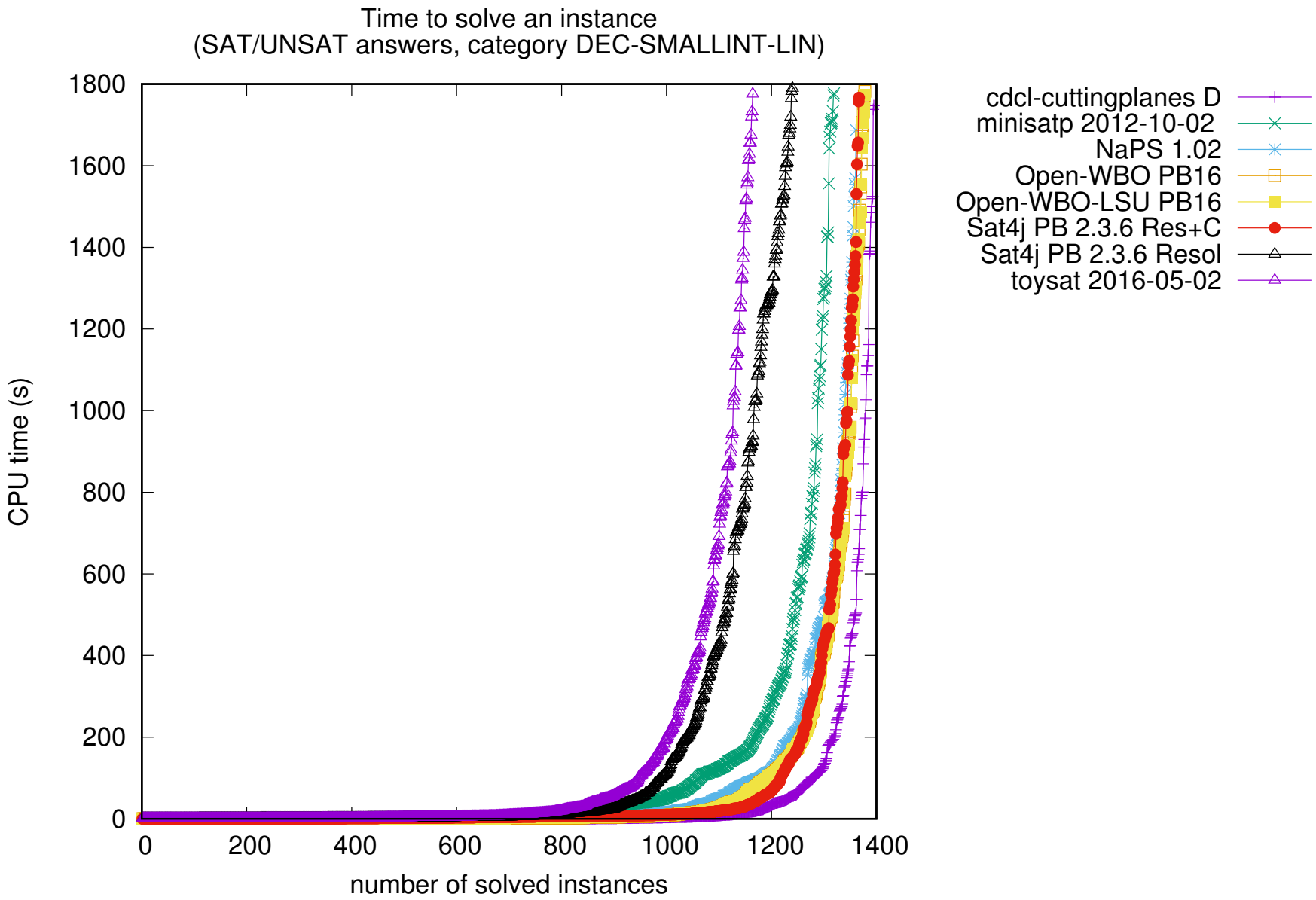
The Virtual Best Solver (VBS)

- is the virtual solver obtained by combining the best results of all submitted solvers.
- **can be obtained by running in parallel all submitted solvers**
- represents the current state of the art (SOTA)
- is a reference for the evaluation of the other solvers

Results for DEC-SMALLINT-LIN

Rank	Solver	#solved	Detail	%inst.	%VBS
Total number of instances: 1783					
Virtual Best Solver (VBS)		1560	402 S, 1158 U	87%	100%
1	cdcl-cp DEC	1395	303 S, 1092 U	78%	89%
2	Open-WBO-LSU	1378	330 S, 1048 U	77%	88%
3	Open-WBO	1378	329 S, 1049 U	77%	88%
4	Sat4j PB Res+CP	1367	315 S, 1052 U	77%	88%
5	NaPS	1361	338 S, 1023 U	76%	87%
6	minisatp 2012	1319	384 S, 935 U	74%	85%
7	Sat4j PB Resolution	1240	342 S, 898 U	70%	79%
8	toysat	1164	323 S, 841 U	65%	75%

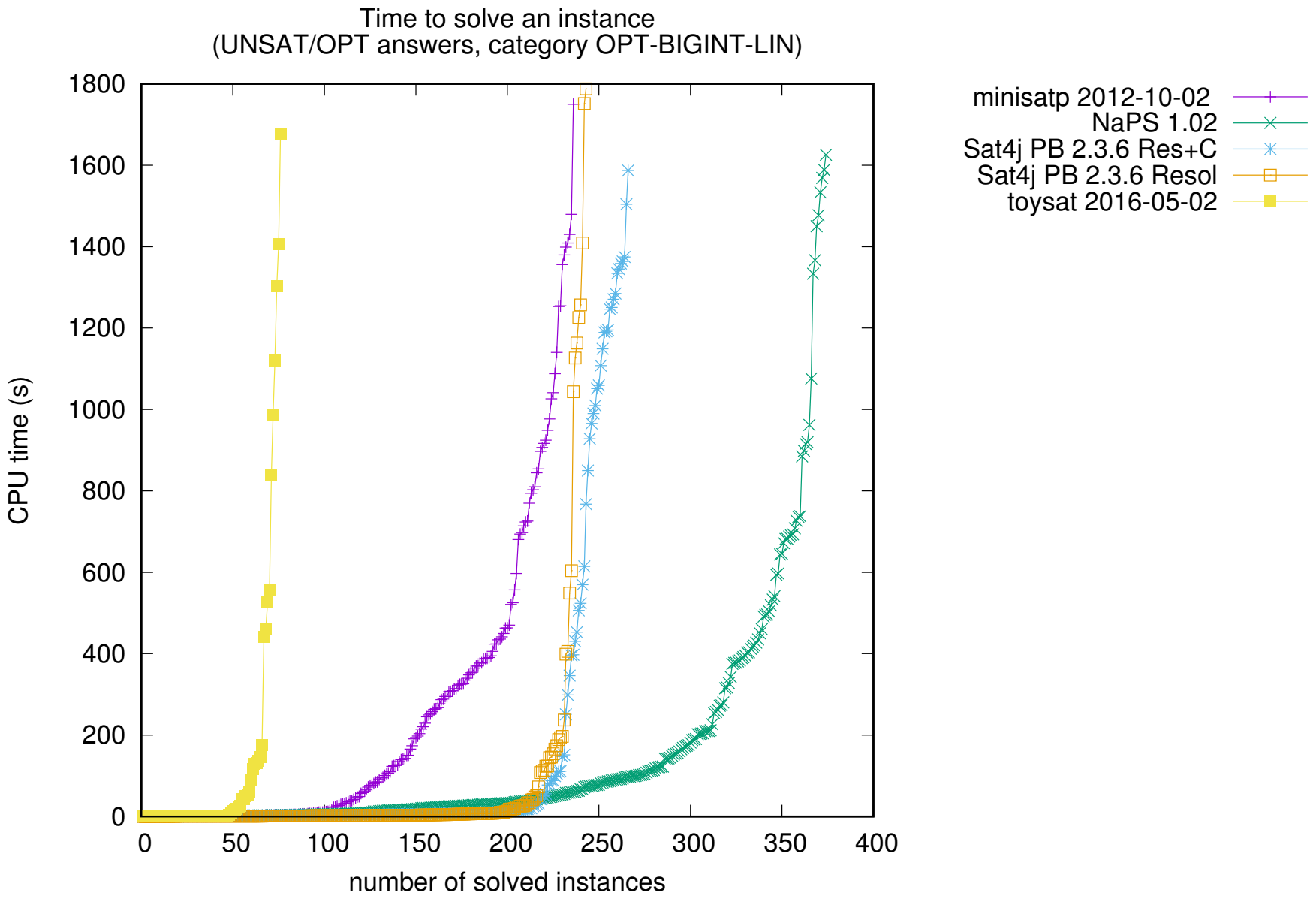
DEC-SMALLINT-LIN



Results for OPT-BIGINT-LIN

Rank	Solver	#solved	Detail	%inst.	%VBS
Total number of instances: 1109					
Virtual Best Solver (VBS)		470	397 O, 73 U	42%	100%
1	NaPS	374	305 O, 69 U	34%	80%
2	Sat4j PB Res+CP	266	196 O, 70 U	24%	57%
3	Sat4j PB Resolution	243	175 O, 68 U	22%	52%
4	minisatp 2012	236	166 O, 70 U	21%	50%
5	toysat	76	38 O, 38 U	7%	16%

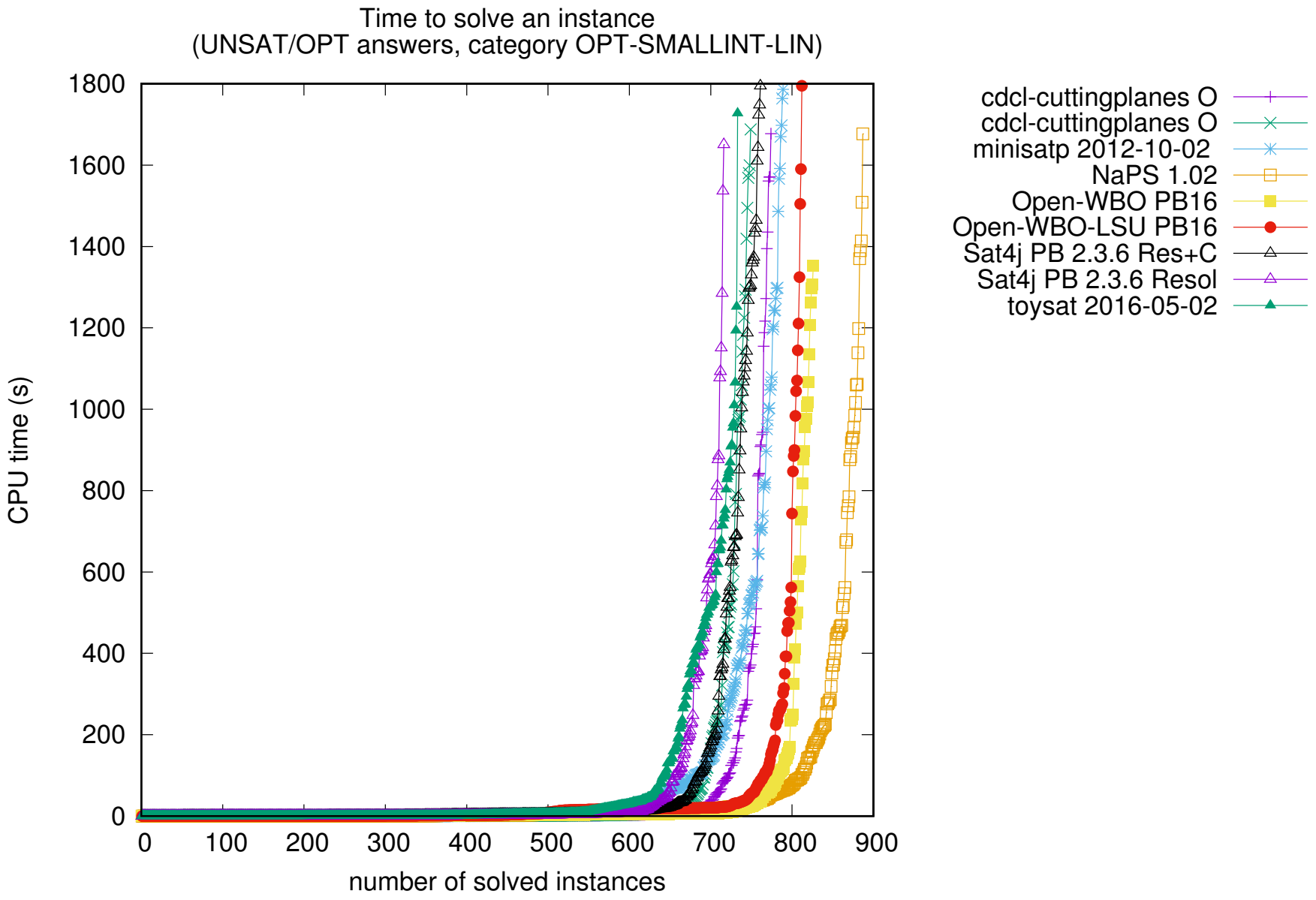
OPT-BIGINT-LIN



Results for OPT-SMALLINT-LIN

Rank	Solver	#solved	Detail	%inst.	%VBS
Total number of instances: 1600					
Virtual Best Solver (VBS)		983	884 O, 99 U	61%	100%
1	NaPS	887	802 O, 85 U	55%	90%
2	Open-WBO	826	781 O, 45 U	52%	84%
3	Open-WBO-LSU	812	767 O, 45 U	51%	83%
4	minisatp 2012	789	713 O, 76 U	49%	80%
5	cdcl-cp OPT bin. search	774	685 O, 89 U	48%	79%
6	Sat4j PB Res+CP	761	672 O, 89 U	48%	77%
7	cdcl-cp OPT lin. search	749	660 O, 89 U	47%	76%
8	toysat	733	670 O, 63 U	46%	75%
9	Sat4j PB Resolution	716	649 O, 67 U	45%	73%

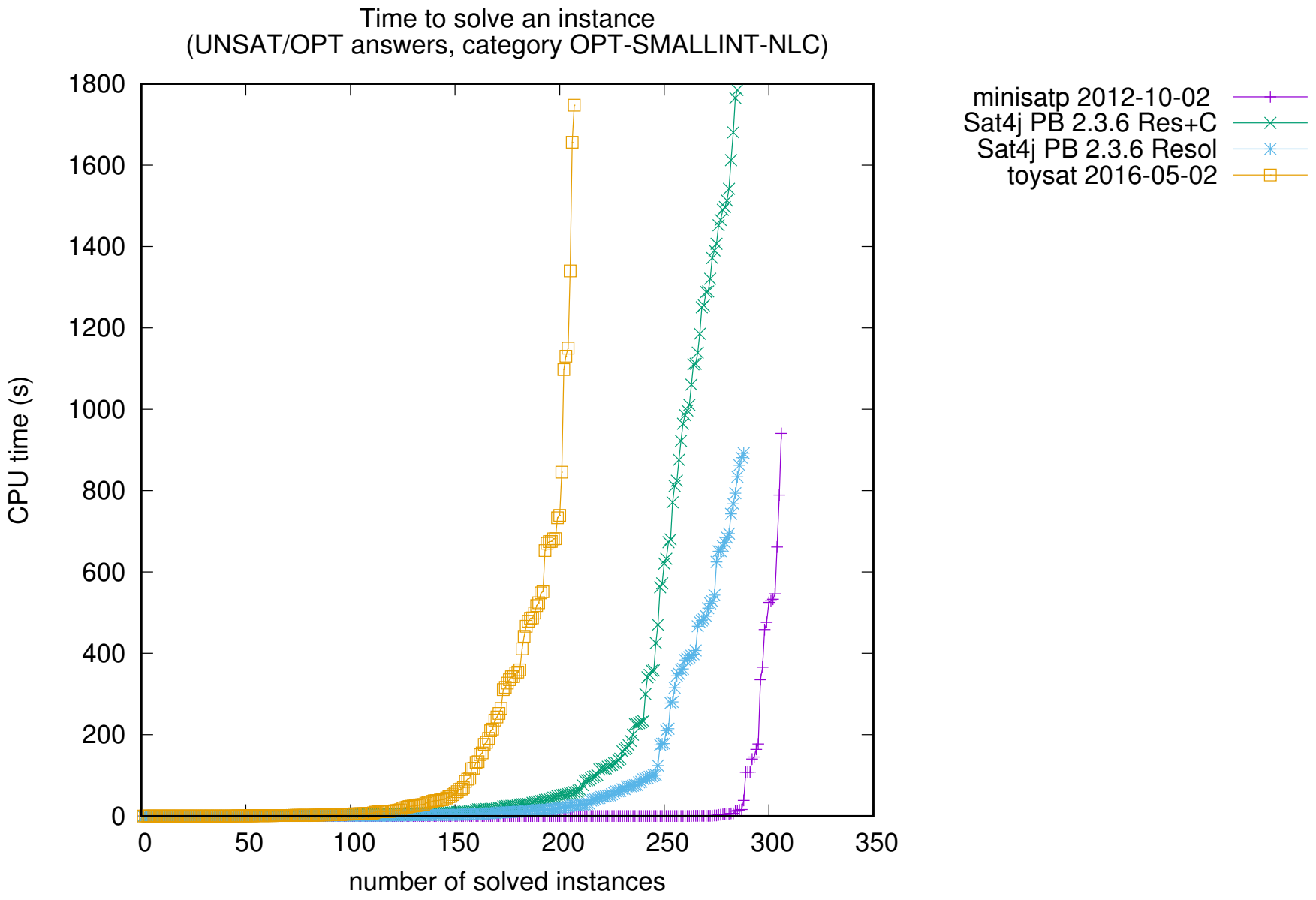
OPT-SMALLINT-LIN



Results for OPT-SMALLINT-NLC

Rank	Solver	#solved	Detail	%inst.	%VBS
Total number of instances: 506					
Virtual Best Solver (VBS)		309	309 O	61%	100%
1	minisatp 2012	306	306 O	60%	99%
2	Sat4j PB Resolution	288	288 O	57%	93%
3	Sat4j PB Res+CP	285	285 O	56%	92%
4	toysat	207	207 O	41%	67%

OPT-SMALLINT-NLC

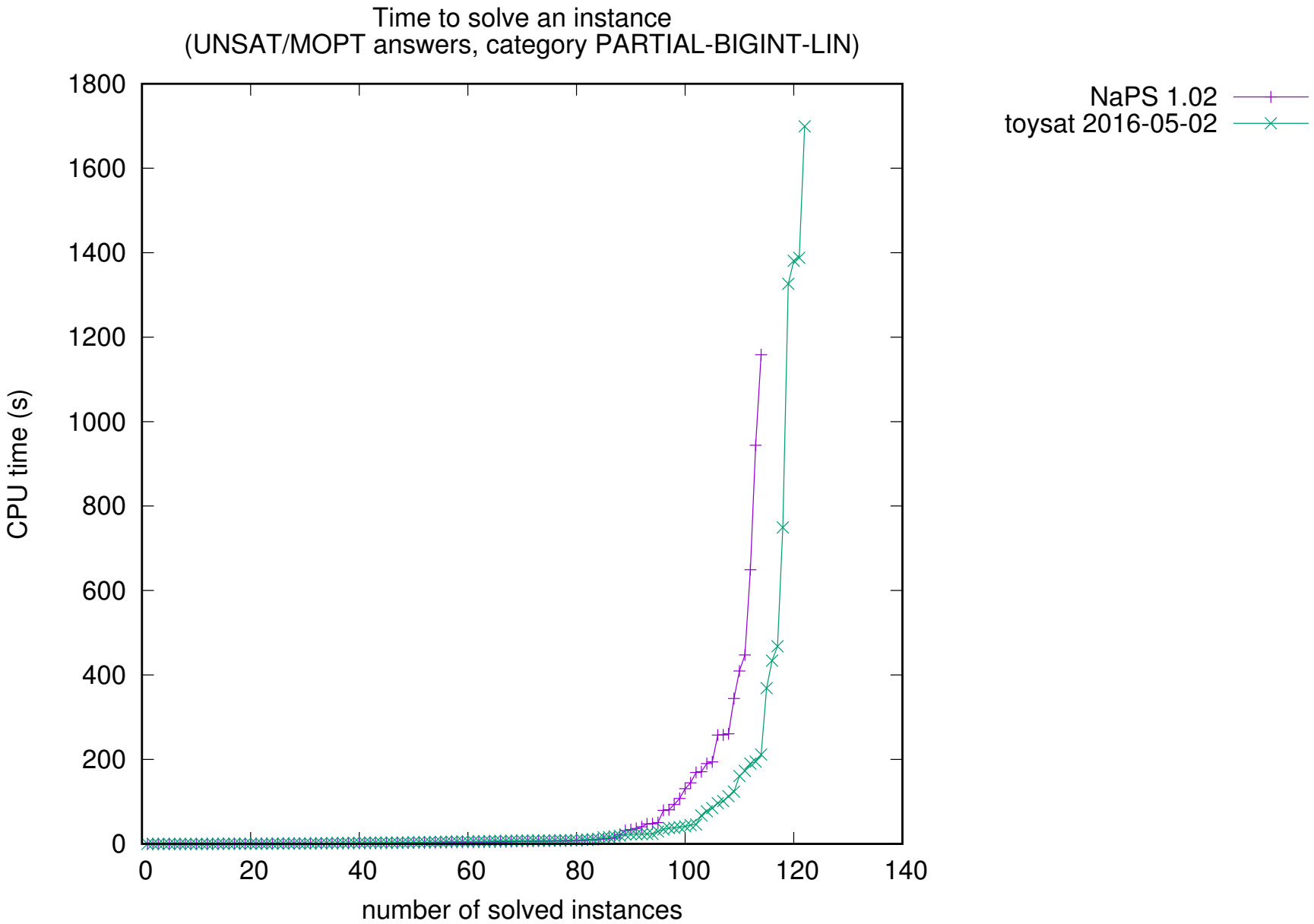


Results for PARTIAL-BIGINT-LIN

Rank	Solver	#solved	Detail	%inst.	%VBS
Total number of instances: 263					
Virtual Best Solver (VBS)		142	142 O	54%	100%
1	toysat	122	122 O	46%	86%
2	NaPS	114	114 O	43%	80%

Sat4J is incorrect in this category (seems to ignore the top cost).

PARTIAL-BIGINT-LIN

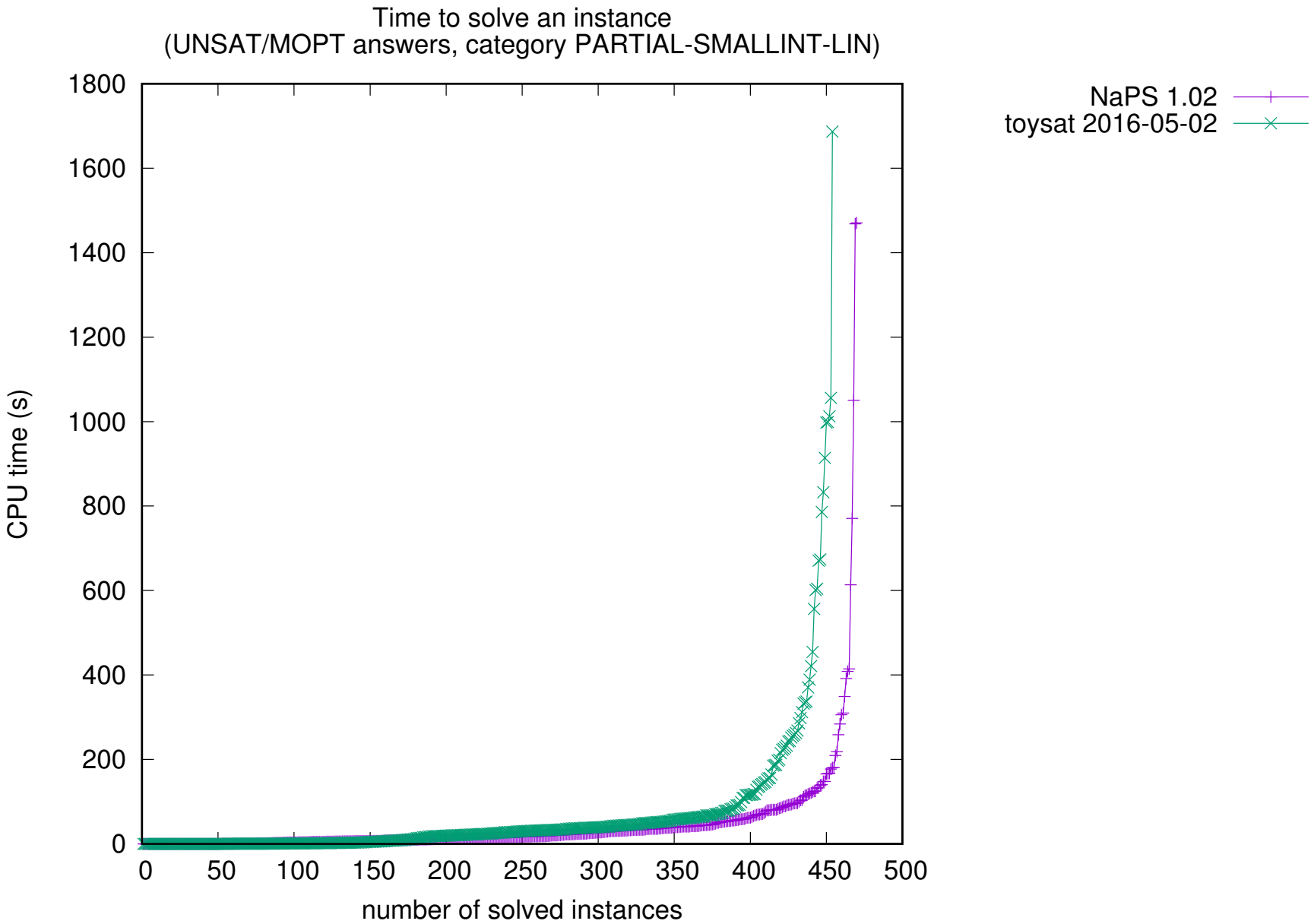


Results for PARTIAL-SMALLINT-LIN

Rank	Solver	#solved	Detail	%inst.	%VBS
Total number of instances: 549					
Virtual Best Solver (VBS)		472	471 O, 1 U	86%	100%
1	NaPS	470	469 O, 1 U	86%	100%
2	toysat	454	453 O, 1 U	83%	96%

Sat4J is incorrect in this category (seems to ignore the top cost).

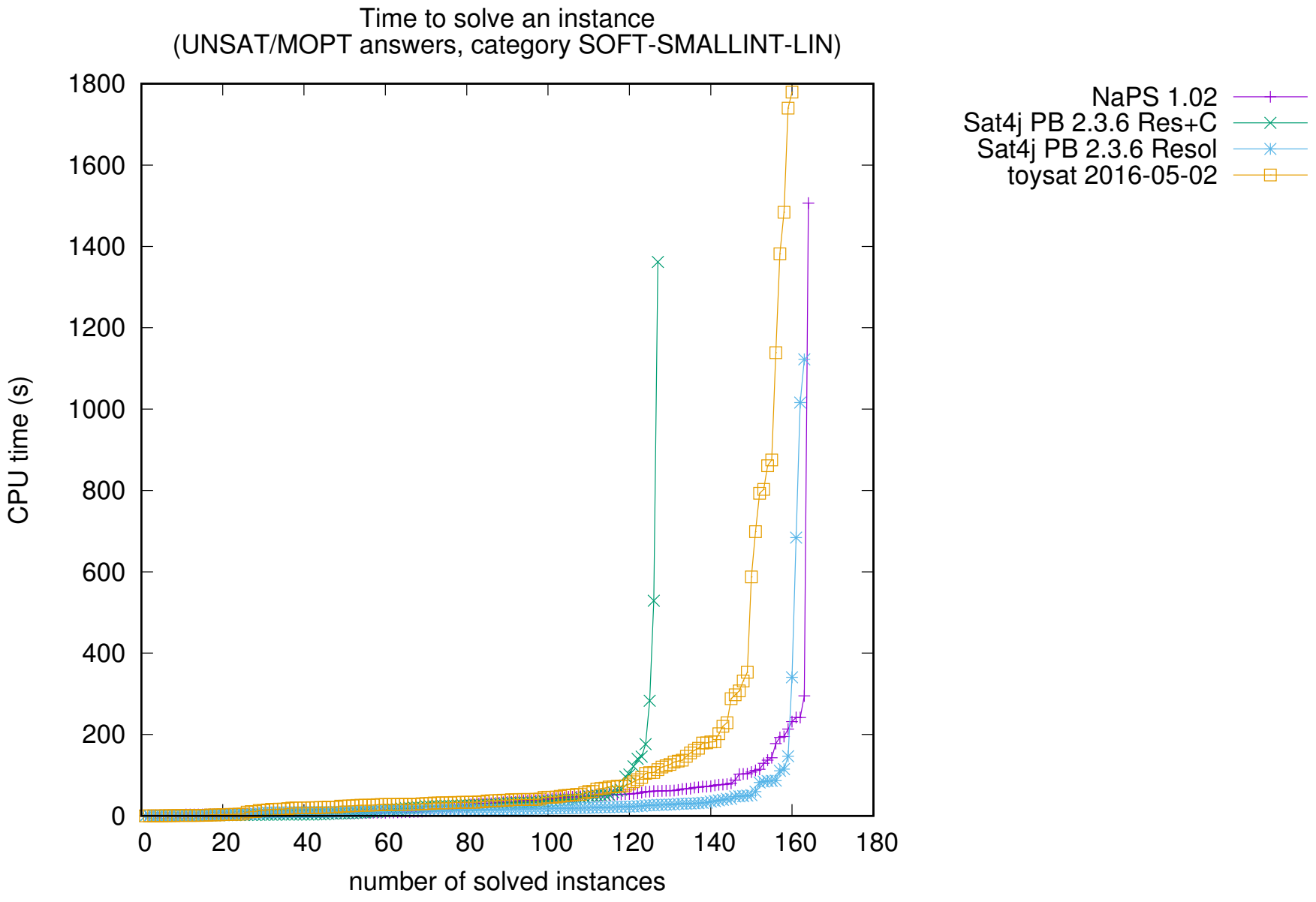
PARTIAL-SMALLINT-LIN



Results for SOFT-SMALLINT-LIN

Rank	Solver	#solved	Detail	%inst.	%VBS
Total number of instances: 201					
Virtual Best Solver (VBS)		195	195 O	97%	100%
1	NaPS	164	164 O	82%	84%
2	Sat4j PB Res	163	163 O	81%	84%
3	toysat	160	160 O	80%	82%
4	Sat4j PB Res+CP	127	127 O	63%	65%

SOFT-SMALLINT-LIN



More information

- All details are on the web site **http://www.cril.univ-artois.fr/PB16/**
- We expect to have a new website in the next few months, allowing better comparison of solvers.
- Thanks to all participants!
- There will be a PB competition or evaluation next year.
- Please consider submitting solvers or instances.