

The Results of SAT Competition 2023

A New Hope
... Strikes Back
The Return of ...

Tomáš Balyo, **Marijn J. H. Heule**, Markus Iser,
Matti Järvisalo, and Martin Suda

SAT 2023 Conference, Alghero (Italy)

July 7, 2023

Competition Overview

SAT Competitions

- ▶ 3 competitions in the 90s (1992,1993, 1996)
- ▶ 15 SAT Competitions (2002–)
- ▶ 5 SAT Races (2006, 2008, 2010, 2015, 2019)
- ▶ 1 SAT Challenge (2012)

Goals

- ▶ Promotion of SAT solvers and their development
- ▶ Compilation of new challenging benchmarks
- ▶ Evaluation of current state-of-the-art solvers

Key rules

- ▶ Certified results of unsatisfiability using proof logging
 - ▶ Instance is “not solved” if proof checker finds times out
- ▶ Disqualification of buggy solvers
 - ▶ Producing an incorrect model
 - ▶ Report UNSAT on a known satisfiable instance
- ▶ Mandatory solver descriptions + open source
- ▶ Ranking scheme: PAR-2
 - ▶ Favors solvers that are faster (not only count solved instances)
- ▶ BYOB (Bring Your Own Benchmarks)
 - ▶ At most 20 instances per participant are used

Competition Summary

Main Track

- ▶ 400 benchmarks
 - ▶ 306 new submissions
 - ▶ 94 of previous competitions
- ▶ 49 sequential solvers
- ▶ 15 parallel solvers
- ▶ 3 cloud solvers

New: Multiple Verified Checkers

Participants picked one of these options:

- ▶ Verified LRAT and LPR Proof Checking with cake_lpr
Yong Kiam Tan, Marijn J. H. Heule, and Magnus O. Myreen
- ▶ GRAT: a formally verified (UN)SAT proof checker
Peter Lammich
- ▶ VeriPB and CakePB: Verified Pseudo-Boolean Proofs
by Bart Bogaerts, Ciaran McCreesh, Magnus O. Myreen,
Jakob Nordström, Andy Oertel, and Yong Kiam Tan

Timeout:

- ▶ Solver: 5000 seconds
- ▶ Checker tool chain: 45000 seconds

A Personal Note

Part of the organizing team since 2013

Rule since 2014: organizers cannot participate

This year three groups at CMU wanted to participate

- ▶ Andrew Haberlandt and Harrison Green
- ▶ Joseph Reeves and Randy Bryant
- ▶ Md Solimul Chowdhury

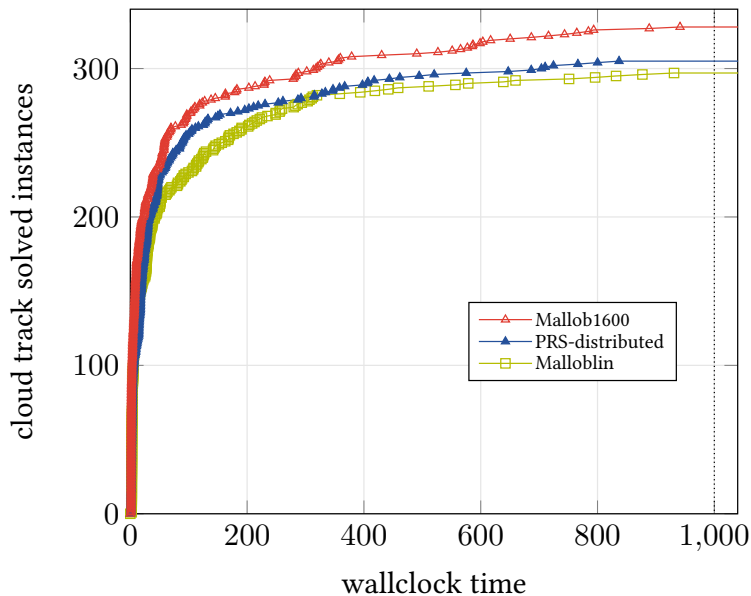
Involved with all of them (but I did not submit). I decided to:

- ▶ Submit no benchmarks
- ▶ Do not participate in benchmark selection
- ▶ Recuse myself from any decisions that impact winners

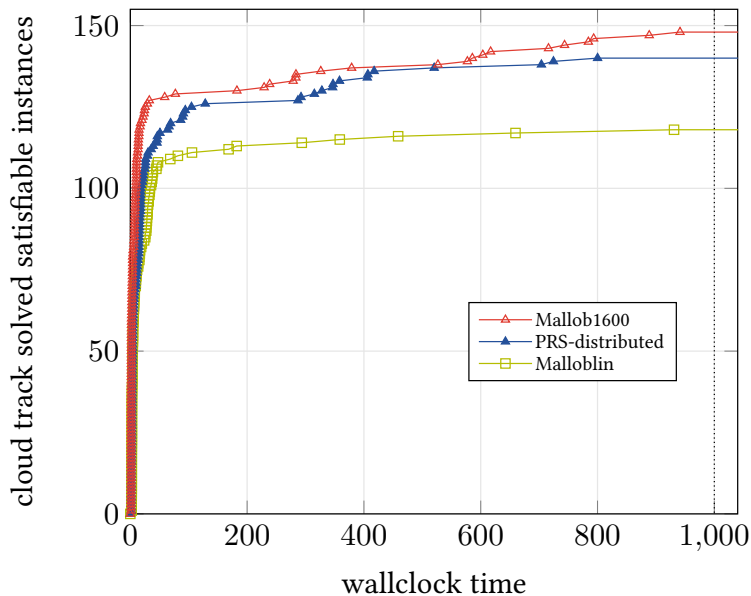
Cloud Track

	Solver	Authors	PAR-2	Solved
1	Mallob1600	Dominik Schreiber	426.10	328
2	PRS-distributed	Zhihan Chen, Xindi Zhang, Yuhang Qianand, and Shaowei Cai	530.92	305
3	Malloblin	Md Solimul Chowdhury	573.53	297

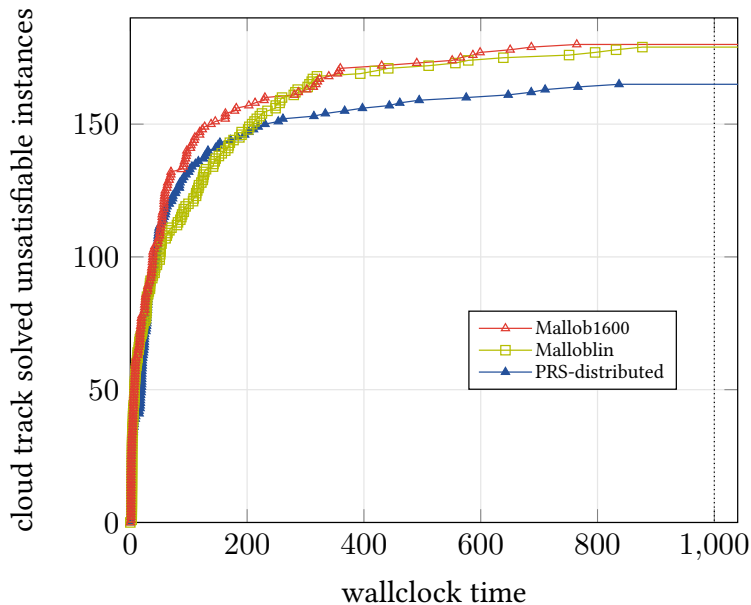
Cloud Track ALL Plot



Cloud Track SAT Plot



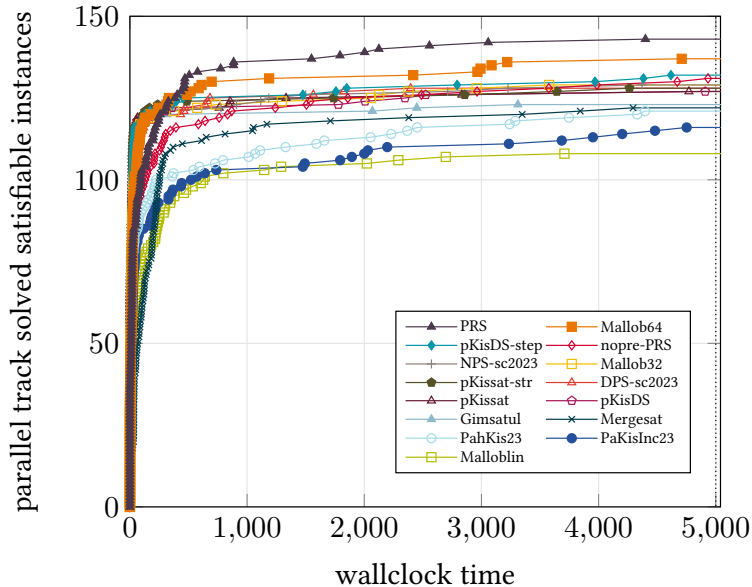
Cloud Track UNSAT Plot



Parallel Track SAT

	Solver	Authors	PAR-2	Solved
1	PRS-parallel	Zhihan Chen, Xindi Zhang, Yuhang Qian and Shaowei Cai	1143.56	143
2	Mallob64	Dominik Schreiber	1505.21	137
3	pKisDS-step	Zhihui Xie, Xu Liu, Wanqian Luo, Junhua Huang, Hui-Ling Zhen, Xijun Li, Mingxuan Yuan and Shuai Li	1796.61	132

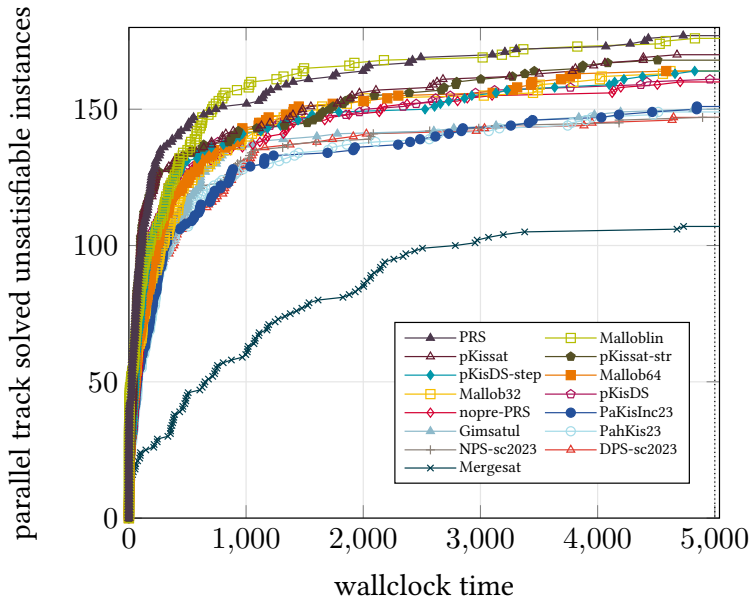
Parallel Track SAT Plot



Parallel Track UNSAT

	Solver	Authors	PAR-2	Solved
1	PRS-parallel	Zhihan Chen, Xindi Zhang, Yuhang Qian and Shaowei Cai	1455.86	177
2	Malloblin	Md Solimul Chowdhury	1501.29	176
3	p-Kissat	Vincent Vallade, Julien Sopena and Souheib Baarir	1852.87	170

Parallel Track UNSAT Plot

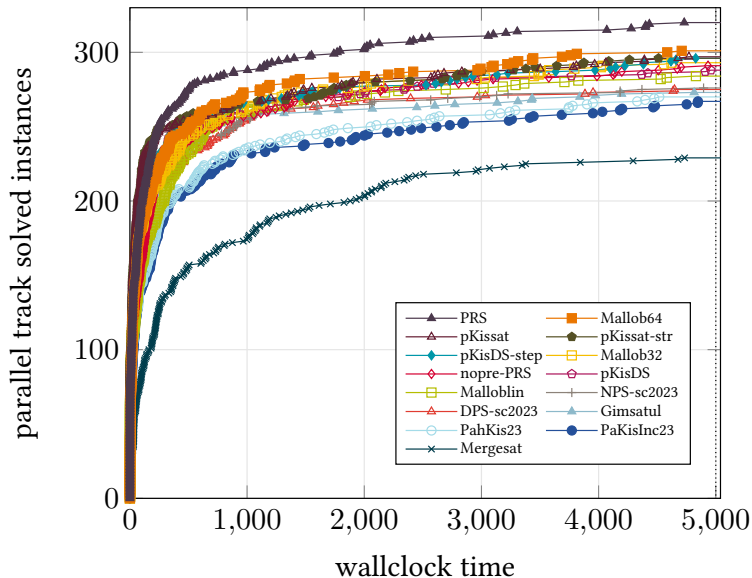


Parallel Track

	Solver	Authors	PAR-2	Solved
1	PRS-parallel	Zhihan Chen, Xindi Zhang, Yuhang Qian and Shaowei Cai	2272.36	320
2	Mallob64	Dominik Schreiber	2746.30	301
3	p-Kissat	Vincent Vallade, Julien Sopena and Souheib Baarir	2824.57	297

- ▶ Similar number solved by parallel and cloud track winner (64 cores and 5000 seconds vs 1600 cores and 1000 seconds)

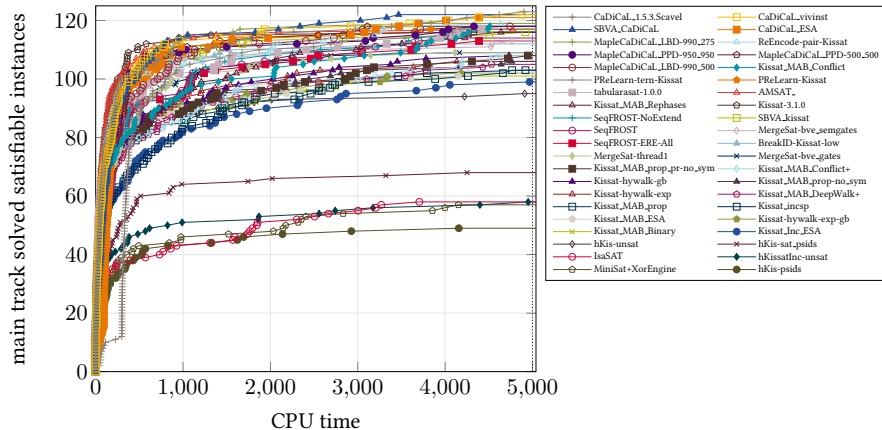
Parallel Track ALL Plot



Main (Sequential) Track SAT

	Solver	Authors	PAR-2	Solved
1	SBVA-CaDiCaL	Andrew Haberlandt and Harrison Green	2337.38	123
2	CaDiCaL-Scavel	Zhihui Li, Guanfeng Wu, Yang Xu, Keming Wang, Zhiguo Long and Zhibin Yu	2351.34	122
3	CaDiCaL-vivinst	Armin Biere, Mathias Fleury and Florian Pollitt	2359.31	122

Main (Sequential) Track SAT Plot

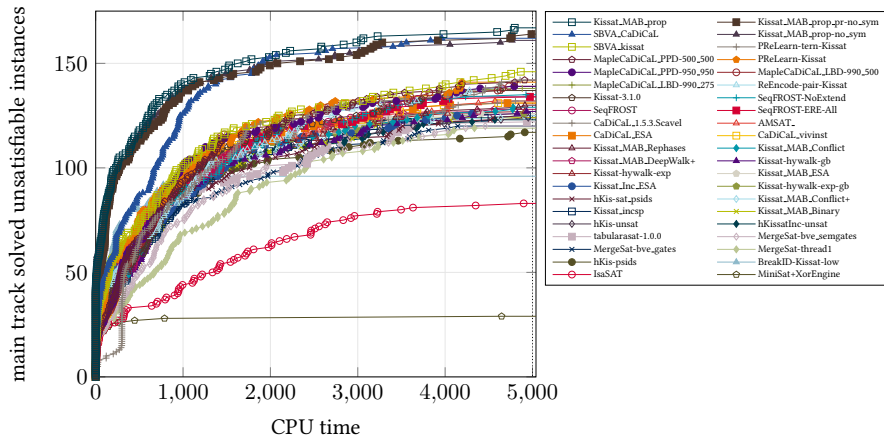


► The Return of CaDiCaL (strongest on satisfiables)

Main Sequential Track UNSAT

	Solver	Authors	PAR-2	Solved
1	Kissat-MAB-prop	Yu Gao	2657.13	167
2	SBVA-CaDiCaL	Andrew Haberlandt and Harrison Green	2975.11	162
3	PReLearn-Kissat	Joseph Reeves and Randy Bryant	4016.81	142

Main (Sequential) Track UNSAT Plot

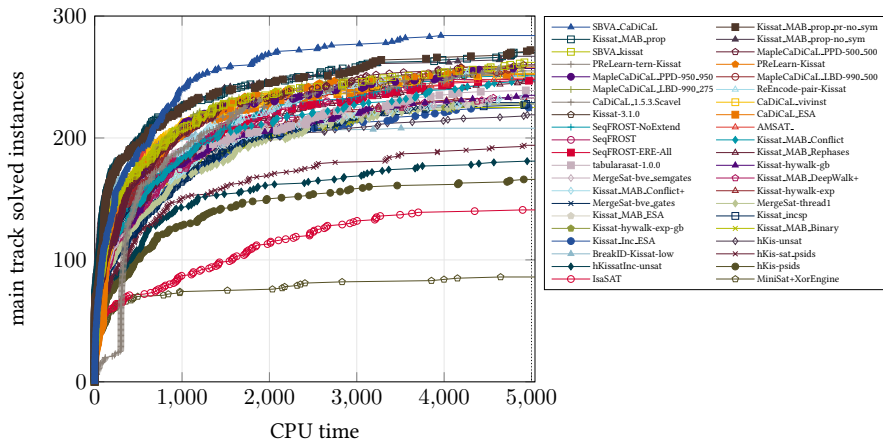


► Beyond Resolution Strikes Back (in all top-tier solvers)

Main Sequential Track

	Solver	Authors	PAR-2	Solved
1	SBVA-CaDiCaL	Andrew Haberlandt and Harrison Green	3274.01	284
2	Kissat-MAB-prop*	Yu Gao	3596.73	272
3	MCaDiCaL-PPD-500	Jonathan Chung, Sam Buss and Vijay Ganesh	3933.51	260

Main (Sequential) Track ALL Plot



► Winner not impacted by own benchmarks

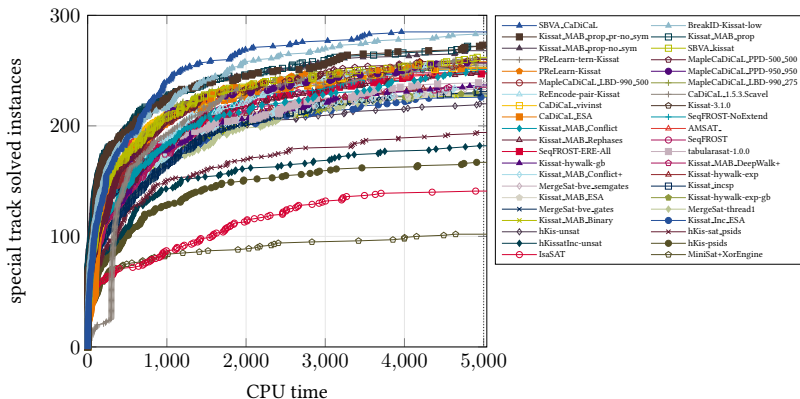
CaDiCaL Hack Track

Winner:

CaDiCaL-vivinst by Armin Biere, Mathias Fleury and Florian Pollitt

Special Prize

What is the ranking if checker timeouts were counted as solved?



Special prize: BreakID-kissat by Bart Bogaerts, Jakob Nordström, Andy Oertel and Çağrı Uluç Yıldırımoglu

Certificate Issues in 2023

- ▶ Rules do not cover corner cases, esp. for UNSAT certificates
 - ▶ Proof checker failure (wrong proof/checker issue)
 - ▶ Proof checker timeouts (45k seconds)
- ▶ New checkers not mature yet
- ▶ *This will change for 2024*

Result	Certificate	SAT	UNSAT (2023)	UNSAT (2024)
Wrong	Wrong	Disq.	Disq.	Disq.
Correct	Timeout	Disq.	Unsolved	Unsolved
Correct	Wrong	Disq.	Unsolved	Disq.

Thank you!

