

The Results of SAT Competition 2024

Marijn Heule, **Markus Iser**, Matti Järvisalo, Martin Suda

SAT 2024 Conference, Pune (India)

August 24, 2024

Competition Anniversary

Anniversary of SAT Competitions

- 3 competitions in the 90s (1992,1993, 1996)
- 16 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
 - ▶ Available Checkers: GRAT, DPR-trim, VeriPB
 - ▶ Instance is “not solved” if proof checker times out
- Disqualification of buggy solvers
 - ▶ Producing an incorrect model or an incorrect proof
- 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

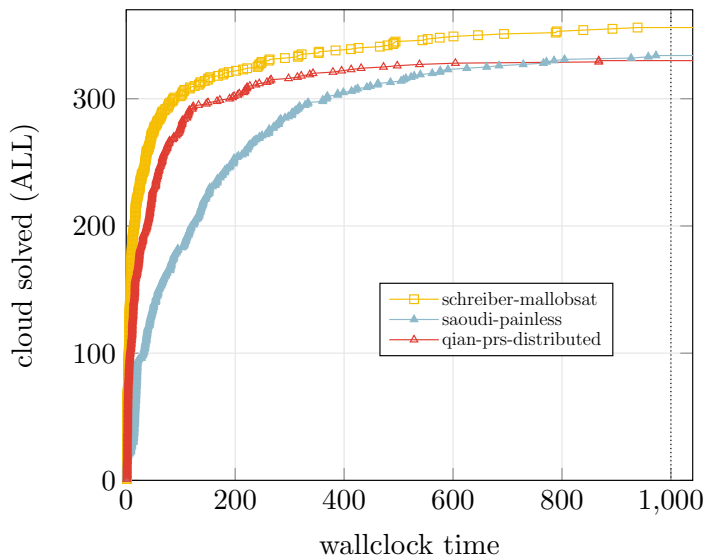
- 300 Previously Unused Instances
 - ▶ 195 Submission from 2024
 - ▶ 31 Submissions from 2023
 - ▶ 74 Submissions from 2022
- 100 Random Instances from Anniversary Track 2022
- 33 Solvers and Configurations
 - ▶ 14 Solvers in Main Track (+7 in No-Limits Track)
 - ▶ 8 Solvers in Parallel Track
 - ▶ 3 Solvers in Cloud Track

Cloud Track

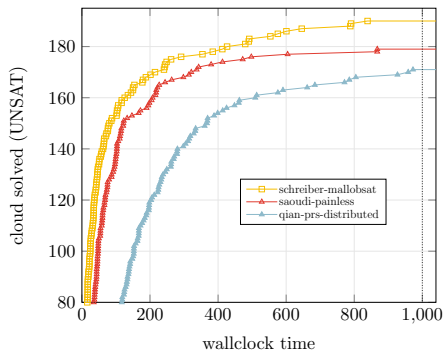
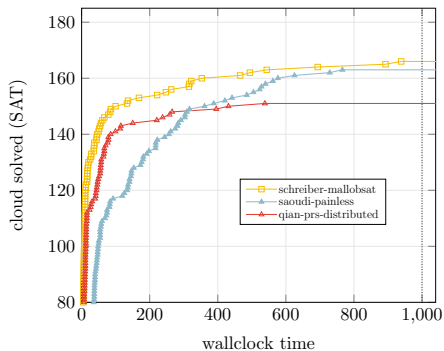
Winning Solvers

	Solver	Authors	PAR-2	Solved
1	MallobSat	Dominik Schreiber	278.87	356
2	PRS-distributed	Yuhang Qian, Zhihan Chen, Xindi Zhang, Shaowei Cai	400.36	330
3	Painless	Mazigh Saoudi, Souheib Baarir, Julien Sopena, Thibault Lejemble, Sabrine Saouli	451.97	334

Cloud Track



Cloud Track SAT and UNSAT

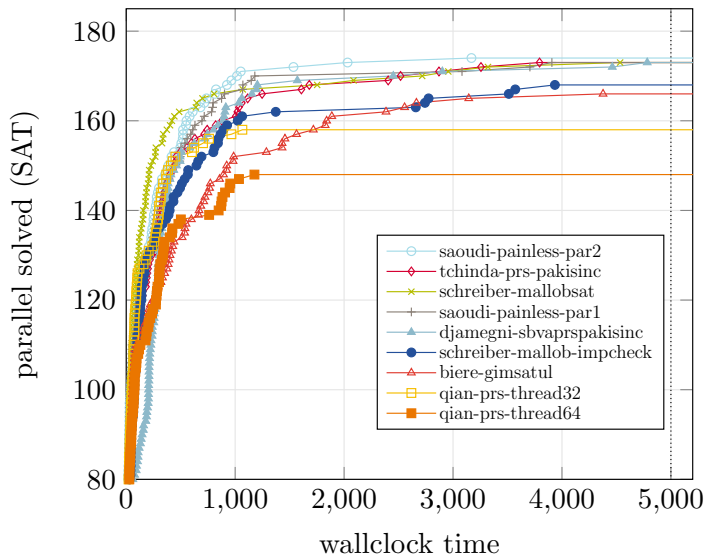


Parallel Track SAT

Winning Solvers

	Solver	Authors	PAR-2	Solved
1	Painless	Mazigh Saoudi, Souheib Baarir, Julien Sopena, Thibault Lejemble, Sabrine Saouli	447.99	174
2	MallobSat	Dominik Schreiber	512.15	173
3	PRS PaKisInc	Rodrigue Konan Tchinda, Clémentin Tayou Djamegni	577.51	173

Parallel Track SAT

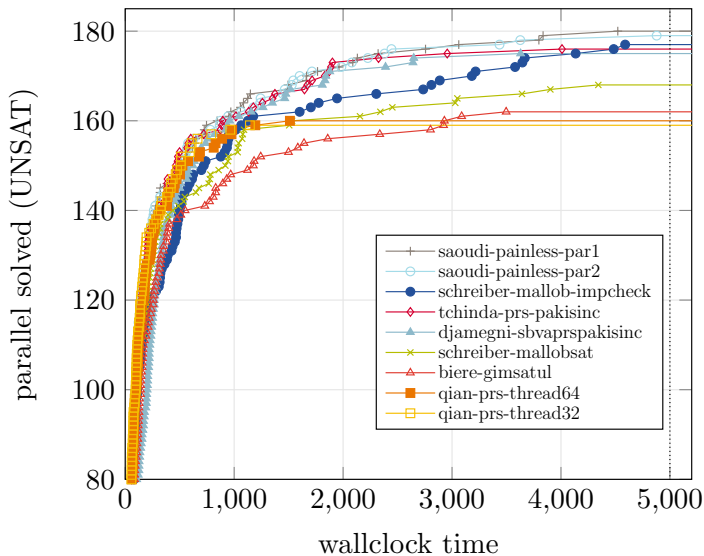


Parallel Track UNSAT

Winning Solvers

	Solver	Authors	PAR-2	Solved
1	Painless	Mazigh Saoudi, Souheib Baarir, Julien Sopena, Thibault Lejemble, Sabrine Saouli	897.63	180
2	PRS PaKisInc	Rodrigue Konan Tchinda, Clémentin Tayou Djamegni	1045.58	176
3	Mallob ImpCheck	Dominik Schreiber	1153.79	177

Parallel Track UNSAT

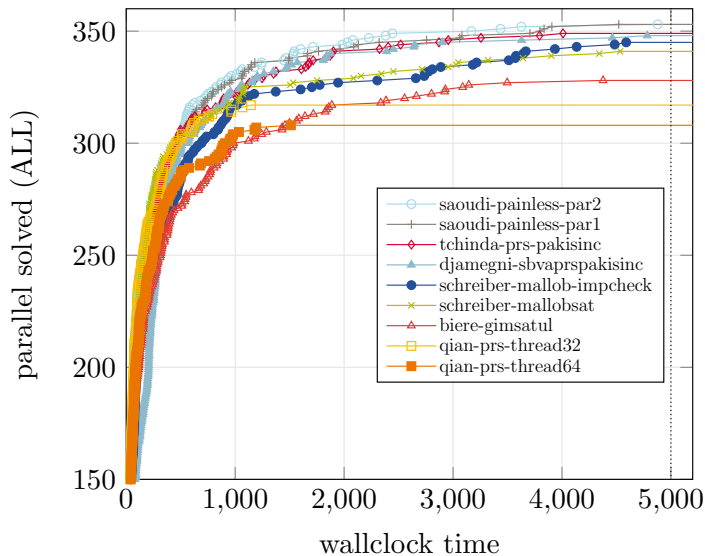


Parallel Track

Winning Solvers

	Solver	Authors	PAR-2	Solved
1	Painless	Mazigh Saoudi, Souheib Baarir, Julien Sopena, Thibault Lejemble, Sabrine Saouli	1397.76	353
2	PRS PaKisInc	Rodrigue Konan Tchinda, Clémentin Tayou Djamegni	1507.70	349
3	Mallob ImpCheck	Dominik Schreiber	1685.25	345

Parallel Track

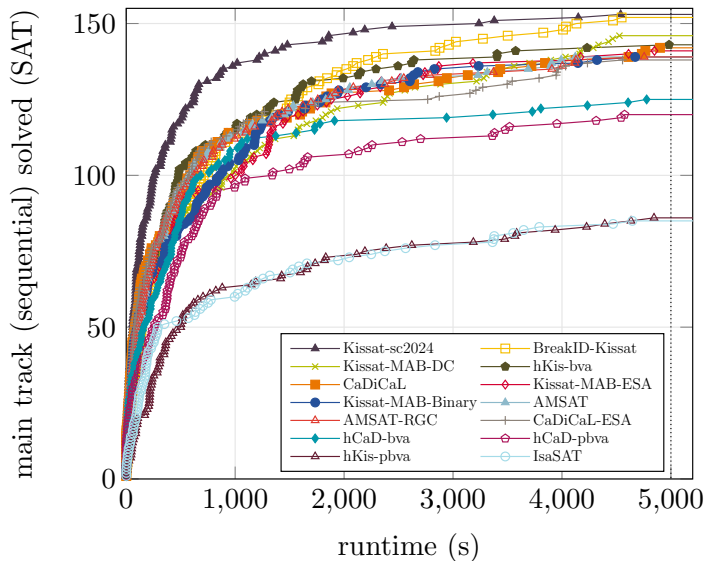


Main Track SAT

Winning Solvers

	Solver	Authors	PAR-2	Solved
1	kissat-sc2024	Armin Biere, Tobias Faller, Katalin Fazekas, Mathias Fleury, Nils Froleys, Florian Pollitt	1270.53	153
2	BreakID-Kissat	Bart Bogaerts, Jakob Nordström, Andy Oertel, Dieter Vandesande	1668.40	152
3	hKis-bva	Rodrigue Konan Tchinda, Clémentin Tayou Djamegni	1997.14	143

Main Track SAT

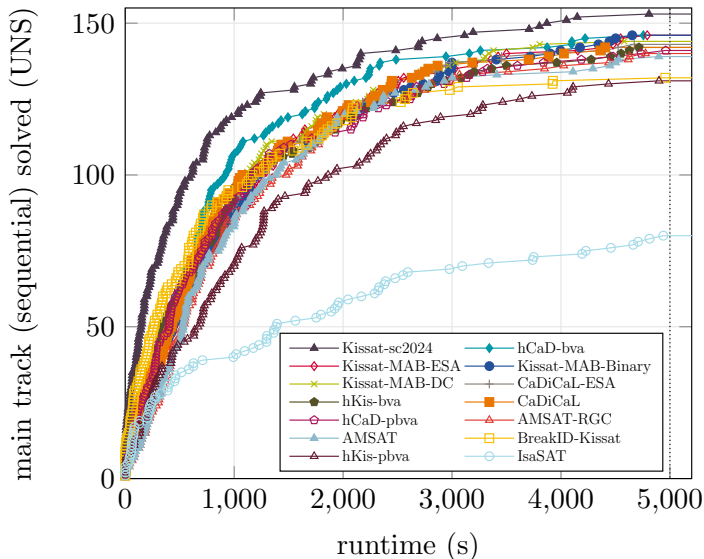


Main Track UNSAT

Winning Solvers

	Solver	Authors	PAR-2	Solved
1	kissat-sc2024	Armin Biere, Tobias Faller, Katalin Fazekas, Mathias Fleury, Nils Froylyks, Florian Pollitt	2077.11	153
2	hCaD-bva	Rodrigue Konan Tchinda, Clémentin Tayou Djamegni	2570.12	146
3	Kissat_MAB-DC	Jinjin Liu, Jianmin Zhang, Yan Sun	2740.35	144

Main Track UNSAT

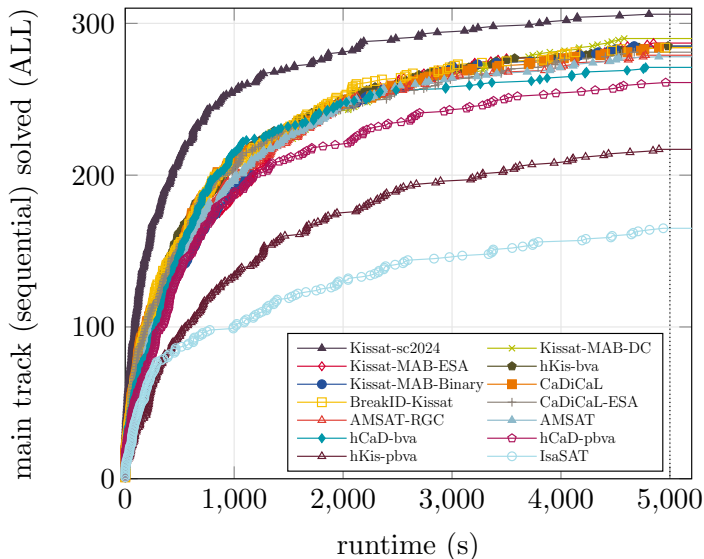


Main Track

Winning Solvers

	Solver	Authors	PAR-2	Solved
1	kissat-sc2024	Armin Biere, Tobias Faller, Katalin Fazekas, Mathias Fleury, Nils Froyeks, Florian Pollitt	2788.13	306
2	Kissat_MAB-DC	Jinjin Liu, Jianmin Zhang, Yan Sun	3435.23	290
3	hKis-bva	Rodrigue Konan Tchinda, Clémentin Tayou Djamegni	3461.15	285

Main Track



Certificates

- Verified checking of models with GRAT in “sat mode”
- For unsatisfiability, we used GRAT, DPR-trim, and VeriPB
- No wrong certificates → found no buggy solvers
- Some could not be checked in time
 - ▶ GRAT: 2 proof-checker timeouts (on same instance)
 - ▶ DPR: 17 proof-checker timeouts (on ten instances)
 - ▶ VeriPB: 25 proof-checker timeouts (on 25 instances)

Issues and Future Competitions

- Some solvers submitted to the No-Limits track produced correct proofs and their source code was available. Why No-Limits?
- New rules for benchmark submissions:
 - ▶ No submission of isomorphic instances. We will provide an isomorphism checker.
 - ▶ Precisely clarify the origin of instances if they are not originally from the submitter. Be kind: Contact the original author.
- Smaller Timeouts for Parallel and Cloud Tracks

Feedback from the Engineers and Scientists at AWS

- Reduce timeouts in the Cloud/Parallel tracks to 200/1000 seconds
 - ▶ results are mostly unchanged
 - ▶ timeouts too long for most applications of stakeholders at AWS
- Implementation and testing quality of submissions has been low
 - ▶ most engineering effort is spent in manually debugging submissions
 - ▶ small fraction of the submissions follow the specified interface of returning UNKNOWN for anything that is not SAT/UNSAT
 - ▶ please clean up: leaving intermediate state around in the solver containers corrupts the next job
- solvers tend to perform best with one solver per physical CPU core
- automation for yes/no testing: increases the burden on submitters
- partner with academic groups and students to work on improving the automation and debugging experience for submitters

Farewell Starexec and Rest in Peace

Starexec was an internationally shared compute cluster of over 275 compute nodes, which ...

- was used by 16 automated reasoning communities
- hosted over 60 competitions
- was for 10 years funded by NSF, across two CRI grants
- had over 1,300 registered users, and 20-40 unique logins per week
- had between 500,000 and several million solver invocations per week

This was the last SAT Competition hosted on Starexec.

- cluster for and form of coming competitions still to be determined
- intermediate, scaled down edition in the spirit of SAT Race in 2025

Recognitions and Thanks

- Marijn Heule, Markus Iser, Matti Järvisalo, Martin Suda (Organizers)
- Robert Jones, Benjamin Jones, Mike Whalen (Amazon)
- Aaron Stump (Starexec)
- Tomáš Balyo (Filuta)

Thank you!

