



# Reviving the Workshop Series on Testing Database Systems – DBTest

Carsten Binnig<sup>1</sup> · Alexander Böhm<sup>2</sup> · Tilmann Rabl<sup>3</sup> · Pinar Tözün<sup>4</sup>

Published online: 1 December 2022

© Gesellschaft für Informatik e.V. and Springer-Verlag GmbH Germany, part of Springer Nature 2022

## Abstract

With the ever increasing complexity of database systems and their pervasive use in industry, testing them has been an important issue for a long time. Recognizing this relevance, researchers and industry have started the Workshop Series on Testing Database Systems in 2008 collocated with ACM SIGMOD. Six instances of the workshop were successfully run until 2013. Five years later, in 2018, we revived the workshop in a new, biannual format. Today, the DBTest workshop consistently has high-quality submissions, expert presenters, and active participants across both academia and industry. Going forward, we plan to open the workshop up to an even more diverse audience, especially the research communities that focus on software testing and debugging in general, and not only on database systems.

**Keywords** DBTest · SIGMOD

## 1 Introduction

Software testing has been an essential topic in computer systems in general. Database systems are complex software systems with large-scale adoption. They have been evolving at a fast pace following the diverse needs of emerging data-intensive applications. The correctness of their behavior and the efficiency and stability of their performance is therefore at high importance. Recognizing this importance, researchers in database community as well as practitioners from industry started the Workshop Series on Testing Database Systems in 2008 collocated with ACM SIGMOD Conference. The workshop ran very successfully every year collocated with SIGMOD from 2008 until 2013. After a five

year hiatus, in 2018, we revived the workshop in a new, biannual format.

Today, the DBTest workshop<sup>1</sup> is a venue for industry and academia to come together and exchange best practices from real-world applications and latest research insights covering the broad topics of testing, debugging, and benchmarking of data-related systems. In addition, the workshop often offers stories of how an academic research prototype can grow into a full-fledged real-world product (e.g., Hyper [7, 8], DuckDB [11]) from the people who build such systems. This way, DBTest also serves as an educational venue for especially junior researchers and developers.

## 2 Focus of the Workshop

As for the topics to focus on during the workshop, we believe that there is the need and interest for discussing topics that are related to not only testing database management systems but also data-intensive systems in general. Specifically, emerging new technologies such as non-volatile memory impose new challenges (e.g., avoiding persistent memory leaks and partial writes), and novel system designs including FPGAs, GPUs, and RDMA call for additional attention and sophistication. Moreover, there are new classes of data-driven applications (i.e., machine learning and big-data scenarios) that need to be considered. Another dimension

---

Carsten Binnig  
carsten.binnig@cs.tu-darmstadt.de

Alexander Böhm  
alexanderboehm@google.com

✉ Tilmann Rabl  
tilmann.rabl@hpi.de

Pinar Tözün  
pito@itu.dk

<sup>1</sup> Darmstadt, Germany

<sup>2</sup> Google, ABC-Str. 19, Hamburg, Germany

<sup>3</sup> Hasso Plattner Institute, Uni Potsdam, Potsdam, Germany

<sup>4</sup> IT University of Copenhagen, Copenhagen, Denmark

<sup>1</sup> <https://dbtest.io/>.

includes completely new system designs – such as crowd-sourcing applications, where testing imposes various monetary and functional challenges (i.e., result set quality and verification), scalable machine-learning systems that combine distributed computing with modern hardware, or even cloud-based data management systems with high release cadence and extreme global reach. Finally, there is an ever-increasing popularity and proliferation of NoSQL systems. These systems are taking new approaches and design decisions (e.g., by replacing the traditional ACID guarantees with relaxed consistency models such as BASE) to increase performance and scalability. Consequentially, they require special testing efforts and rigor to ensure classical database strengths such as reliability, integrity, and performance can be successfully carried over to these novel system architectures and designs.

Following these focus areas, the workshop accepts submissions and presentations on the set of topics listed below:

- Testing of database systems, storage services, and database applications
- Testing of database systems using novel hardware and software technology (non-volatile memory, hardware transactional memory, ...)
- Testing heterogeneous systems with hardware accelerators (GPUs, FPGAs, ASICs, ...)
- Testing distributed and big-data systems
- Testing machine learning systems
- Specific challenges of testing and quality assurance for cloud-based systems
- War stories and lessons learned
- Performance and scalability testing
- Testing the reliability and availability of database systems
- Algorithms and techniques for automatic program verification
- Maximizing code coverage when testing database systems and applications
- Generation of synthetic data for test databases
- Testing the effectiveness of adaptive policies and components
- Tools for analyzing database management systems (profilers, debuggers, ...)
- Workload characterization with respect to performance metrics and engine components
- Metrics for test quality, robustness, efficiency, and effectiveness
- Operational aspects such as continuous integration and delivery pipelines
- Security and vulnerability testing

In the rest of this paper, we first cover the highlights from the different instances of DBTest. Then, we summarize the plans for the future of the DBTest workshop.

### 3 From 2008 to 2013

The first DBTest workshop was collocated with SIGMOD 2008 in Vancouver, Canada [5]. It was organized by Leo Giakoumakis and Donald Kossmann. DBTest 2008 had a clear focus on testing with a significant number of contributions from academia and industry. DBTest 2009, which was collocated with SIGMOD 2009 in Providence, USA, was organized by Carsten Binnig and Benoit Dageville [3]. Again, the workshop featured many contributions by academia and industry, mostly on the topic of testing, but also workload modelling and benchmarking. In 2010, DBTest was organized by Shivnath Babu and Glenn Paulley, collocated with SIGMOD in Indianapolis, USA [1]. For the first time, DBTest also included a panel with experts from industry and academia, discussing the grand challenges in database system testing. DBTest 2011, which was collocated with SIGMOD in Athens, Greece, was organized by Goetz Graefe and Kenneth Salem [6]. The 2011 instance of DBTest saw a stronger shift towards tooling and benchmarks. The trend of tooling continued in DBTest 2012, which was held in Scottsdale, USA, and organized by Eric Lo and Florian Waas [9]. DBTest 2012 introduced auto-tuning into the program and saw submissions on the topic of big data. The 2013 issue of DBTest, collocated with SIGMOD in New York, USA, was focused on data generation and tooling for database analysis [10]. DBTest 2013 was organized by Vivek Narasayya and Neoklis Polyzotis. DBTest 2013 was the last issue after a series of six yearly workshops.

### 4 From 2018 to 2022

After four years of absence, DBTest was successfully *re-vived* at SIGMOD 2018 in Houston, USA, with over 40 participants, 15 paper submissions (out of which 8 were accepted), and three industry sponsors [2]. DBTest 2018 was organized by Alexander Böhm and Tilmann Rabl as chairs together with Carsten Binnig who was involved in the early iterations of the workshop<sup>2</sup>. In the reincarnation of 2018, there was a lot of positive feedback and we were encouraged by the audience to re-establish DBTest in a yearly cadence. Hence, we decided to only run with a biannual

<sup>2</sup> Today, Alexander Böhm, Tilmann Rabl and Carsten Binnig form the steering committee of the workshop.

cadence to prevent from running out of topics and high-quality submissions.

DBTest 2020 was then organized by Pinar Tözün and Alexander Böhm. With the global challenges caused by the COVID-19 pandemic, both SIGMOD and DBTest moved to an online format in 2020 [14]. Despite these challenges, the interest in the DBTest workshop increased further, with now more than 150 registrations, and over 70 online participants during sessions. With seven accepted papers, three keynotes and four industry sponsors and a diverse set of topics ranging from debugging query compilers, to automated performance testing, to correctness tests for big-data systems, the workshop was well received by the audience.

The last instance of DBTest took place in hybrid mode co-located with SIGMOD 2022 conference in Philadelphia, USA and was organized by Manuel Rigger and Pinar Tözün [13]. The program consisted of one panel, two keynotes, and four paper presentations with two industry sponsors. It covered topics ranging from software testing practices in large-scale projects, automatic bug finding, benchmark generation for systems with different characteristics. 15–35 people participated in-person during sessions, while 15–25 participated online. The sessions with most participation was, as in previous years, the keynote sessions. Both the in-person and virtual audience actively participated in all Q&A sessions highlighting the overall interest in the workshop topics.

One of the remarkable aspects of each DBTest instance has been the high number of industry participants and submissions creating an interesting mix of both academic work and industry perspectives. The workshop has always created interesting lively discussions among the academic and industry participants. Such exchanges enrich DBTest to also be an educational platform for everyone when it comes to building realistic and practical benchmarking and testings infrastructures.

## 5 Looking into the Future

DBTest is a workshop at the intersection of database and software engineering research. Therefore, we are currently reaching out to the software engineering community to see if an alternating collocation with a software engineering conference equivalent to SIGMOD would be sensible and possible.

Techniques like fuzzing have been a topic in DBTest from the start (e.g., Garcia [4]) and continue to be relevant (e.g., Rehman et al. [12]). These have their roots in the software engineering community and debugging and testing database systems is a relevant application.

## 6 Conclusion

In the article, we briefly summarized the history of the International Workshop on Database Testing (DBTest), which currently is collocated with SIGMOD in a biyearly cadence. DBTest is a venue for research in testing, debugging, and benchmarking of database systems and related technologies. All papers are available through the ACM digital library and listed in DBLP<sup>3</sup>.

## Declarations

**Conflict of interest** The authors do not have any financial or non-financial interests that are directly or indirectly related to the work submitted for publication.

## References

1. Babu S, Paulley GN (eds) (2010) Proceedings of the third international workshop on testing database systems. Indianapolis, 7 June 2010. DBTest 2010. ACM, New York
2. Böhm A, Rabl T (eds) (2018) Proceedings of the 7th international workshop on testing database systems. DBTest@SIGMOD 2018, Houston, 15 June 2018 ACM, New York
3. Dageville B, Binnig C (eds) (2009) Proceedings of the 2nd international workshop on testing database systems. DBTest 2009, Providence, Rhode Island, 29 June 2009 ACM, New York
4. Garcia R (2009) Case study: experiences on SQL language fuzz testing. In: Dageville B, Binnig C (eds) Proceedings of the 2nd International Workshop on Testing Database Systems DBTest 2009, Providence, 29 June 2009 ACM, New York
5. Giakoumakis L, Kossmann D (eds) (2008) Proceedings of the 1st international workshop on testing database systems. DBTest 2008, Vancouver, 13 June 2008 ACM, New York
6. Graefe G, Salem K (eds) (2011) Proceedings of the fourth international workshop on testing database systems. DBTest 2011, Athens, 13 June 2011 ACM, New York
7. Kemper A, Neumann T, Funke F, Leis V, Mühe H (2012) Hyper: adapting columnar main-memory data management for transactional AND query processing. *IEEE Data Eng Bull* 35(1):46–51
8. Kersten T, Neumann T (2020) On another level: how to debug compiling query engines. In: Tözün P, Böhm A (eds) Proceedings of the 8th international workshop on testing database systems DBTest@SIGMOD 2020, Portland, 19 June 2020. vol 2. ACM, New York, pp 1–2
9. Lo E, Waas F (eds) (2012) Proceedings of the fifth international workshop on testing database systems. DBTest 2012, Scottsdale, 21 May 2012 ACM, New York
10. Narasayya VR, Polyzotis N (eds) (2013) Proceedings of the sixth international workshop on testing database systems. DBTest 2013, New York, 24 June 2013 ACM, New York, NY, USA
11. Raasveldt M, Mühleisen H (2019) Duckdb: an embeddable analytical database. In: Boncz PA, Manegold S, Ailamaki A, Deshpande A, Kraska T (eds) Proceedings of the 2019 International Conference on Management of Data SIGMOD Conference 2019, Amsterdam, June 30–July 5, 2019 ACM, New York, pp 1981–1984
12. Rehman MS, Elmore AJ (2022) Fuzzydata: a scalable workload generator for testing dataframe workflow systems. In: Rigger M,

<sup>3</sup> DBLP page – International Workshop on Testing Database Systems – <https://dblp.uni-trier.de/db/conf/dbtest-ws/index.html>.

- Tözün P (eds) DBTest@SIGMOD '22: Proceedings of the 9th International Workshop of Testing Database Systems Philadelphia, 17 June 2022 ACM, New York, pp 17–24
13. Rigger M, Tözün P (eds) (2022) DBtest@SIGMOD '22: proceedings of the 9th international workshop of testing database systems. Philadelphia, 17 June 2022 ACM, New York
  14. Tözün P, Böhm A (eds) (2020) Proceedings of the 8th international workshop on testing database systems. DBTest@SIGMOD 2020, Portland, 19 June 2020 ACM, New York