Marcus **Paradies**

marcus.paradies@tu-ilmenau.de | https://marcusparadies.github.io/ | @marcus_paradies

Education

Ph.D. *Dresden, Germany*

Technische Universität Dresden *Jun. 2011 - Feb. 2017*

- Thesis title: Graph Processing in Main-Memory Column Stores
- Advisor: Prof. Wolfgang Lehner

Diploma (Dipl.-Inf.) *Ilmenau, Germany*

Ilmenau University of Technology *Oct. 2005 - Nov. 2011*

- Thesis title: MAXIM: Massive-Scale Entity Matching across XML Data Sources in MapReduce
- Advisor: Prof. Kai-Uwe Sattler

Professional Experience

Technische Universität Ilmenau *Ilmenau, Germany*

PostDoc Researcher *Jul. 2023 - today*

- Research, mentoring, and acquisition of third-party funding in the area of distributed data management.

German Aerospace Center (DLR) *Jena, Germany*

Acting director of Institute of Data Science *Jul. 2022 - Jun. 2023*

- Research, management, mentoring, and acquisition of third-party funding in the area of Data Science.
- Management of 44 scientific and 10 administrative employees

German Aerospace Center (DLR) *Jena, Germany*

Department head Data management and enrichment *Mar. 2022 - Jun. 2023*

- Research, management, mentoring, and acquisition of third-party funding in the area of distributed data management.
- Scientific advisor + management of 4 research groups with 16 employees

German Aerospace Center (DLR) *Jena, Germany*

Acting department head Data management and enrichment *Jan. 2022 - Feb. 2022*

- Research, management, mentoring, and acquisition of third-party funding in the area of distributed data management.
- Scientific advisor + management of 4 research groups with 16 employees

German Aerospace Center (DLR) *Jena, Germany*

Acting department head Data management and analysis *May 2020 - Dec. 2021*

- Research, management, mentoring, and acquisition of third-party funding in the area of distributed data management.
- Scientific advisor + management of 4 research groups with 20 employees

German Aerospace Center (DLR) *Jena, Germany*

Group lead (Scientific Advisor) *Mar. 2018 - Apr. 2020*

- Research, mentoring, and acquisition of third-party funding in the area of distributed data management.
- Scientific advisor of 7 internal and 2 external employees

SAP SE *Walldorf, Germany*

Senior software developer & architect *Jan. 2017 - Feb. 2018*

- Architect and lead developer of SAP HANA Graph and GraphScript.
- Design and implementation of the GraphScript domain-specific language, compiler, and runtime

SAP SE *Walldorf, Germany*

Software Developer *Sep. 2015 - Dec. 2016*

- Architect and lead developer of SAP HANA Graph.
- Design and implementation of SAP HANA Graph index structures and APIs

SAP SE *Walldorf, Germany*

Researcher *Jun. 2011 - Aug. 2015*

- Research on in-memory graph processing in SAP HANA.
- Design and implementation of column store extensions for graph data, traversal operators, index structures, and graph DSL compile + runtime

IBM Watson T.J. Research Center *Hawthorne, US*

Research Internship *Apr. 2010 - Mar. 2011*

- Research in the area of distributed entity matching on semistructured data.

Teaching

WS 2023	**Lecture**, Advanced Database Systems (3 SWS), TU Ilmenau	*Ilmenau*
SS 2023	**Lecture**, Data Storage Systems (3 SWS), TU Ilmenau	*Ilmenau*
SS 2022	**Lecture**, Data Storage Systems (3 SWS), TU Ilmenau	*Ilmenau*
WS 2021	**Lecture**, Database Systems 1 (4 SWS), FSU Jena	*Jena*
SS 2021	**Lecture**, Data Storage Systems (2 SWS), FSU Jena	*Jena*
2018	**Guest lecture**, Seminar *Data Science*, FSU Jena	*Jena*
2014	**Guest lecture**, Seminar *Graph Data Management*, TU Dresden	*Dresden*
2010 - 2011	**Lectures and Hands-On Labs**, IBM DB2 XML Customer Lectures (Storage, Indexing, XQuery)	*US,Canada*

Community Service

Reviewing

2023	**Reviewer**, CHEOPS Workshop@EuroSys
2022	**Reviewer**, CHEOPS Workshop@EuroSys
2021	**Shadow PC Reviewer**, EuroSys
2021	**Reviewer**, GRADES-NDA Workshop@SIGMOD
2021	**Reviewer**, BTW
2020	**Reviewer**, GRADES-NDA Workshop@SIGMOD
2019	**Reviewer**, EURO-PAR
2019	**Reviewer**, GRADES-NDA Workshop@SIGMOD
2018	**Reviewer**, GRADES-NDA Workshop@SIGMOD

Memberships

since 2021	**Member**, TZLR e.V.
since 2018	**Member**, Michael Stifel Center Jena
since 2018	**Member**, Gesellschaft für Informatik e.V.
2016 - 2018	**Member**, LDBC Graph Query Language Task Force
2016 - 2018	**Member**, LDBC SNB Benchmarking Task Force
2017 - 2018	**Member**, ISO Ad-Hoc Group on SQL/Graph, SQL Standards Committee

Organization of Events

2019	**Co-Organizer**, LWDA'19 Track FG-DB
2017	**Organizer**, 9th LDBC TUC Meeting, SAP Headquarters
2016	**Co-Organizer**, 1. German Graph Community Meeting

Patents

2020	**US10546021B2**, Adjacency structures for executing graph algorithms in a relational database
2020	**US10769188B2**, Text searches on graph data
2019	**US10394855B2**, Graph-modeled data processing in a relational database
2019	**US20190303505A1**, Graph data derivatives
2019	**US20190303506A1**, Adaptable adjacency structure for querying graph data
2019	**US20190311060A1**, Extended path finding operations on graph data
2019	**US10503781B2**, Extending graph traversals with application logic
2018	**US9934324B2**, Index structure to accelerate graph traversal
2017	**US9547728B2**, Graph traversal operator and extensible framework inside a column store
2017	**US20170011099A1**, Interactive exploration of large graphs
2015	**US9031976B2**, Flexible tables

Publications

[1] Joshua Reibert, Arne Osterthun, and Marcus Paradies. Meduse: Interactive and visual exploration of ionospheric data. In Birgitta König-Ries, Stefanie Scherzinger, Wolfgang Lehner, and Gottfried Vossen, editors, *Datenbanksysteme für Business, Technologie und Web (BTW 2023), 20. Fachtagung des GI-Fachbereichs „Datenbanken und Informationssysteme" (DBIS), 06.-10, März 2023, Dresden, Germany, Proceedings*, volume P-331 of *LNI*, pages 681–686. Gesellschaft für Informatik e.V., 2023.

[2] Patrick Damme, Marius Birkenbach, Constatinos Bitsakos, Matthias Boehm, Philippe Bonnet, Florina Ciorba, Mark Dokter, Pawel Dowgiallo, Ahmed Eleliemy, Christian Faerber, Georgios Goumas, Dirk Habich, Niclas Hedam, Marlies Hofer, Wenjun Huang, Kevin Innerebner, Vasileios Karakostas, Roman Kern, Tomaž Kosar, Alexander Krause, Daniel Krems, Andreas Laber, Wolfgang Lehner, Eric Mier, Tilmann Rabl, Piotr Ratuszniak, Pedro Silva, Nikolai Skuppin, Andreas Starzacher, Benjamin Steinwender, Ilin Tolovski, Pinar Tözün, Wojciech Ulatowski, Yuanyuan Wang, Izajasz Wrosz, Aleš Zamuda, Ce Zhang, and Xiao Xiang Zhu. DAPHNE: An Open and Extensible System Infrastructure for Integrated Data Analysis Pipelines. In *12th Annual Conference on Innovative Data Systems Research (CIDR '22)*, 2022.

[3] Hani Al-Sayeh, Bunjamin Memishi, Muhammad Attahir Jibril, Marcus Paradies, and Kai-Uwe Sattler. Juggler: Autonomous cost optimization and performance prediction of big data applications. In Zachary Ives, Angela Bonifati, and Amr El Abbadi, editors, *SIGMOD '22: International Conference on Management of Data, Philadelphia, PA, USA, June 12 - 17, 2022*, pages 1840–1854. ACM, 2022.

[4] Michael A. C. Johnson, Marcus Paradies, Marta Dembska, Kristen Lackeos, Hans-Rainer Klöckner, David J. Champion, and Sirko Schindler. Astronomical pipeline provenance: A use case evaluation. In Tanu Malik and Thomas Pasquier, editors, *13th International Workshop on Theory and Practice of Provenance, TaPP 2021, July 19-20, 2021*. USENIX Association, 2021.

[5] Wenjun Huang and Marcus Paradies. An Evaluation of WebAssembly and eBPF as Offloading Mechanisms in the Context of Computational Storage. *CoRR*, abs/2111.01947, 2021.

[6] Renzo Angles, János Benjamin Antal, Alex Averbuch, Peter A. Boncz, Orri Erling, Andrey Gubichev, Vlad Haprian, Moritz Kaufmann, Josep Lluís Larriba-Pey, Norbert Martínez-Bazan, József Marton, Marcus Paradies, Minh-Duc Pham, Arnau Prat-Pérez, Mirko Spasic, Benjamin A. Steer, Gábor Szárnyas, and Jack Waudby. The LDBC social network benchmark. *CoRR*, abs/2001.02299, 2020.

[7] Marcus Paradies. CryoDrill: Near-Data Processing in Deep and Cold Storage Hierarchies. In *CIDR 2019, 9th Biennial Conference on Innovative Data Systems Research*, 2019.

[8] Bunjamin Memishi, Raja Appuswamy, and Marcus Paradies. Cold Storage Data Archives: More Than Just a Bunch of Tapes. In *Proceedings of the 15th International Workshop on Data Management on New Hardware, DaMoN 2019*, pages 1:1–1:7, 2019.

[9] Matthias Hauck, Marcus Paradies, and Holger Fröning. Software-Based Buffering of Associative Operations on Random Memory Addresses. In *2019 IEEE International Parallel and Distributed Processing Symposium, IPDPS*, pages 943–952, 2019.

[10] Marcus Paradies, Stefan Plantikow, and Oskar van Rest. Graph Data Management Systems. In *Encyclopedia of Big Data Technologies*. 2019.

[11] Bunjamin Memishi, Raja Appuswamy, and Marcus Paradies. Cold Storage Data Archives: More Than Just a Bunch of Tapes. *CoRR*, abs/1904.04736, 2019.

[12] Marcus Paradies and Hannes Voigt. Graph Representations and Storage. In *Encyclopedia of Big Data Technologies*. 2019.

[13] Frank Tetzel, Hannes Voigt, Marcus Paradies, Romans Kasperovics, and Wolfgang Lehner. Analysis of Data Structures Involved in RPQ Evaluation. In *Proceedings of the 7th International Conference on Data Science, Technology and Applications, DATA*, pages 334–343, 2018.

[14] Gábor Szárnyas, Arnau Prat-Pérez, Alex Averbuch, József Marton, Marcus Paradies, Moritz Kaufmann, Orri Erling, Peter A. Boncz, Vlad Haprian, and János Benjamin Antal. An early look at the LDBC social network benchmark's business intelligence workload. In *Proceedings of the 1st ACM SIGMOD Joint International Workshop on Graph Data Management Experiences & Systems (GRADES) and Network Data Analytics (NDA)*, pages 9:1–9:11, 2018.

[15] Marcus Paradies, Sirko Schindler, Stephan Kiemle, and Eberhard Mikusch. Large-Scale Data Management for Earth Observation Data - Challenges and Opportunities. In *Proceedings of the Conference "Lernen, Wissen, Daten, Analysen", LWDA*, pages 285–288, 2018.

[16] Renzo Angles, Marcelo Arenas, Pablo Barceló, Peter A. Boncz, George H. L. Fletcher, Claudio Gutierrez, Tobias Lindaaker, Marcus Paradies, Stefan Plantikow, Juan F. Sequeda, Oskar van Rest, and Hannes Voigt. G-CORE: A Core for Future Graph Query Languages. In *Proceedings of the 2018 International Conference on Management of Data, SIGMOD Conference*, pages 1421–1432, 2018.

[17] Hannes Voigt, Marcus Paradies, and Theo Härder. Editorial. *Datenbank-Spektrum*, 17(2):97–99, 2017.

[18] Marcus Paradies and Hannes Voigt. Big Graph Data Analytics on Single Machines - An Overview. *Datenbank-Spektrum*, 17(2):101–112, 2017.

[19] Marcus Paradies, Cornelia Kinder, Jan Bross, Thomas Fischer, Romans Kasperovics, and Hinnerk Gildhoff. GraphScript: implementing complex graph algorithms in SAP HANA. In *Proceedings of The 16th International Symposium on Database Programming Languages, DBPL*, pages 13:1–13:4, 2017.

[20] Matthias Hauck, Marcus Paradies, and Holger Fröning. Can Modern Graph Processing Engines Run Concurrent Queries Efficiently? In *Proceedings of the Fifth International Workshop on Graph Data-management Experiences & Systems, GRADES@SIGMOD/PODS*, pages 5:1–5:6, 2017.

[21] Frank Tetzel, Hannes Voigt, Marcus Paradies, and Wolfgang Lehner. An Analysis of the Feasibility of Graph Compression Techniques for Indexing Regular Path Queries. In *Proceedings of the Fifth International Workshop on Graph Data-management Experiences & Systems, GRADES@SIGMOD/PODS*, pages 11:1–11:6, 2017.

[22] Jan Broß, Simon Gog, Matthias Hauck, and Marcus Paradies. Fast Construction of Compressed Web Graphs. In *String Processing and Information Retrieval - 24th International Symposium, SPIRE*, pages 116–128, 2017.

[23] Renzo Angles, Marcelo Arenas, Pablo Barceló, Peter A. Boncz, George H. L. Fletcher, Claudio Gutierrez, Tobias Lindaaker, Marcus Paradies, Stefan Plantikow, Juan F. Sequeda, Oskar van Rest, and Hannes Voigt. G-CORE: A Core for Future Graph Query Languages. *CoRR*, abs/1712.01550, 2017.

[24] Matthias Hauck, Marcus Paradies, Holger Fröning, Wolfgang Lehner, and Hannes Rauhe. Highspeed Graph Processing Exploiting Main-Memory Column Stores. In *Euro-Par 2015: Parallel Processing Workshops - Euro-Par 2015 International Workshops*, pages 503–514, 2015.

[25] Marcus Paradies, Wolfgang Lehner, and Christof Bornhövd. GRAPHITE: an extensible graph traversal framework for relational database management systems. In *Proceedings of the 27th International Conference on Scientific and Statistical Database Management, SSDBM '15*, pages 29:1–29:12, 2015.

[26] Max Wildemann, Michael Rudolf, and Marcus Paradies. The Time Has Come: Traversal and Reachability in Time-Varying Graphs. In *Biomedical Data Management and Graph Online Querying - VLDB 2015 Workshops, Big-O(Q) and DMAH*, pages 169–183, 2015.

[27] Marcus Paradies, Elena Vasilyeva, Adrian Mocan, and Wolfgang Lehner. Robust Cardinality Estimation for Subgraph Isomorphism Queries on Property Graphs. In *Biomedical Data Management and Graph Online Querying - VLDB 2015 Workshops, Big-O(Q) and DMAH*, pages 184–198, 2015.

[28] Marcus Paradies, Michael Rudolf, and Wolfgang Lehner. GraphVista: Interactive Exploration Of Large Graphs. *CoRR*, abs/1506.00394, 2015.

[29] Marcus Paradies, Michael Rudolf, Christof Bornhövd, and Wolfgang Lehner. GRATIN: Accelerating Graph Traversals in Main-Memory Column Stores. In *Second International Workshop on Graph Data Management Experiences and Systems, GRADES 2014, co-located with SIGMOD/PODS*, pages 9:1–9:6, 2014.

[30] Marcus Paradies, Wolfgang Lehner, and Christof Bornhövd. GRAPHITE: An Extensible Graph Traversal Framework for Relational Database Management Systems. *CoRR*, abs/1412.6477, 2014.

[31] Michael Rudolf, Marcus Paradies, Christof Bornhövd, and Wolfgang Lehner. The Graph Story of the SAP HANA Database. In *Datenbanksysteme für Business, Technologie und Web (BTW), 15. Fachtagung des GI-Fachbereichs "Datenbanken und Informationssysteme" (DBIS)*, pages 403–420, 2013.

[32] Michael Rudolf, Marcus Paradies, Christof Bornhövd, and Wolfgang Lehner. SynopSys: large graph analytics in the SAP HANA database through summarization. In *First International Workshop on Graph Data Management Experiences and Systems, GRADES 2013, co-loated with SIGMOD/PODS*, page 16, 2013.

[33] Marcus Paradies, Susan Malaika, Jérôme Siméon, Shahan Khatchadourian, and Kai-Uwe Sattler. Entity matching for semistructured data in the Cloud. In *Proceedings of the ACM Symposium on Applied Computing, SAC*, pages 453–458, 2012.

[34] Marcus Paradies. An Efficient Blocking Technique for Reference Matching using MapReduce. *Datenbank-Spektrum*, 11(1):47–49, 2011.

[35] Marcus Paradies. An Efficient Blocking Technique for Reference Matching using MapReduce. In *Proceedings BTW 2011 - Workshops und Studierendenprogramm*, pages 110–114, 2011.

[36] Marcus Paradies, Christian Lemke, Hasso Plattner, Wolfgang Lehner, Kai-Uwe Sattler, Alexander Zeier, and Jens Krüger. How to juggle columns: an entropy-based approach for table compression. In *Fourteenth International Database Engineering and Applications Symposium (IDEAS 2010)*, pages 205–215, 2010.

[37] Marcus Paradies, Susan Malaika, Matthias Nicola, and Kevin Xie. Comparing XML processing performance in middleware and database: a case study. In *Proceedings of the 11th International Middleware Conference Industrial Track*, pages 35–39, 2010.