

46th International Conference on Very Large Data Bases (VLDB 2020)

Proceedings of the VLDB Endowment Volume 13

Online

31 August – 4 September 2020

Volume 1 of 5

ISBN: 978-1-7138-1612-6

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2020) by VLDB Endowment

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-ncnd/4.0/>. For any use beyond those covered by this license, obtain permission by emailing info@vldb.org.

Printed with permission by Curran Associates, Inc. (2020)

For permission requests, please contact VLDB Endowment at the address below.

VLDB Endowment
Attn: Marianne Siroker
353 Serra Mall – Stanford University
Stanford CA 94305

Phone: (650) 723-0872

Fax: (650) 725-2588

info@vldb.org.

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

Research Papers

Revenue Maximization for Query Pricing	1
<i>Shuchi Chawla, Shaleen Deep, Paraschos Koutris, Yifeng Teng</i>	
Realtime Top-k Personalized PageRank over Large Graphs on GPUs	15
<i>Jieming SHI, Renchi Yang, Tianyuan J Jin, Xiaokui Xiao, Yin Yang</i>	
Fast Large-Scale Trajectory Clustering	29
<i>Sheng Wang, Zhifeng Bao, J. Shane Culpepper, Timos Sellis, Xiaolin Qin</i>	
Automating Distributed Tiered Storage Management in Cluster Computing.....	43
<i>Herodotos Herodotou, Elena Kakoulli</i>	
APOLLO: Automatic Detection and Diagnosis of Performance Regressions in Database Systems...	57
<i>Jinho Jung, Hong Hu, Joy Arulraj, Taesoo Kim, Woonhak Kang</i>	
Lowering the Latency of Data Processing Pipelines Through FPGA based Hardware Acceleration.	71
<i>Muhsen Owaida, Gustavo Alonso, Laura Fogliarini, Anthony Hock-Koon, Pierre-Etienne Melet</i>	

Research Papers

Model Slicing for Supporting Complex Analytics with Elastic Inference Cost and Resource Constraints	86
<i>Shaofeng Cai, Gang Chen, Beng Chin Ooi, Jinyang Gao</i>	
Cross-chain Deals and Adversarial Commerce	100
<i>Maurice Herlihy, Liuba Shrira, Barbar Liskov</i>	
Efficient Main-Memory Top-K Selection For Multicore Architectures	114
<i>Vasileios Zois, Vassilis Tsotras, Walid Najjar</i>	
Scalable Garbage Collection for In-Memory MVCC Systems	128
<i>Jan Böttcher, Viktor Leis, Thomas Neumann, Alfons Kemper</i>	
Fully Dynamic Depth-First Search in Directed Graphs.....	142
<i>Bohua Yang, Dong Wen, Lu Qin, Ying Zhang, Xubo Wang, Xuemin Lin</i>	
LINC: A Motif Counting Algorithm for Uncertain Graphs	155
<i>Chenhao Ma, Reynold Cheng, Laks V.S. Lakshmanan, Tobias Grubenmann, Yixiang Fang, Xiaodong Li</i>	
ObliDB: Oblivious Query Processing for Secure Databases.....	169
<i>Saba Eskandarian, Matei Zaharia</i>	
Secure Multi-Party Functional Dependency Discovery.....	184
<i>Chang Ge, Ihab Ilyas, Florian Kerschbaum</i>	
Vacuum Filters: More Space-Efficient and Faster Replacement for Bloom and Cuckoo Filters	197
<i>Minmei Wang, Mingxun Zhou, Shouqian Shi, Chen Qian</i>	
On Supporting Efficient Snapshot Isolation for Hybrid Workloads with Multi-Versioned Indexes ...	211
<i>Yihan Sun, Guy Blelloch, Wan Shen Lim, Andrew Pavlo</i>	

Research Papers

Interleaved Multi-Vectorizing	226
<i>Zhuhe Fang, Beilei Zheng, Chuliang Weng</i>	
A Unified Optimization Algorithm For Solving “Regret-Minimizing Representative” Problems.....	239
<i>Suraj Shetiya, Abolfazl Asudeh, Sadia Ahmed, Gautam Das</i>	
Pushing Data-Induced Predicates Through Joins in Big-Data Clusters.....	252
<i>Srikanth Kandula, Laurel Orr, Surajit Chaudhuri</i>	
Discovery of Approximate (and Exact) Denial Constraints.....	266
<i>Eduardo H. M. Pena, Eduardo C. de Almeida, Felix Naumann</i>	
Deep Unsupervised Cardinality Estimation.....	279
<i>Zongheng Yang, Eric Liang, Amog Kamsetty, Chenggang Wu, Yan Duan, Xi Chen, Pieter Abbeel, Joseph M. Hellerstein, Sanjay Krishnan, Ion Stoica</i>	
Free Gap Information from the Differentially Private Sparse Vector and Noisy Max Mechanisms..	293
<i>Zeyu Ding, Yuxin Wang, Danfeng Zhang, Daniel Kifer</i>	
An End-to-End Learning-based Cost Estimator	307
<i>Ji Sun, Guoliang Li</i>	
Last-Mile Delivery Made Practical: An Efficient Route Planning Framework with Theoretical Guarantees	320
<i>Yuxiang Zeng, Yongxin Tong, Lei Chen</i>	
Database Processing-in-Memory: An Experimental Study	334
<i>Tiago R. Kepe, Eduardo C. de Almeida, Marco A. Z. Alves</i>	
Incorporating Super-Operators in Big-Data Query Optimizers	348
<i>Jyoti Leeka, Kaushik Rajan</i>	
Efficient Progressive Minimum k-Core Search	362
<i>Conggai Li, Fan Zhang, Ying Zhang, Lu Qin, Wenjie Zhang, Xuemin Lin</i>	
Harmonia: Near-Linear Scalability for Replicated Storage with In-Network Conflict Detection.....	376
<i>Hang Zhu, Zhihao Bai, Jialin Li, Elliss Michael, Dan R. K. Ports, Ion Stoica, Xin Jin</i>	
Learning to Sample: Counting with Complex Queries	390
<i>Brett Walenz, Stavros Sintos, Sudeepa Roy, Jun Yang</i>	
Return of the Lernaean Hydra: Experimental Evaluation of Data Series Approximate Similarity Search	403
<i>Karima Echihabi, Kostas Zoumpatianos, Themis Palpanas, Houda Benbrahim</i>	

Research Papers

DPTree: Differential Indexing for Persistent Memory.....	421
<i>Xinjing Zhou, Lidan Shou, Ke Chen, Wei Hu, Gang Chen</i>	
AJoin: Ad-hoc Stream Joins at Scale.....	435
<i>Suraj Shetiya, Abolfazl Asudeh, Sadia Ahmed, Gautam Das</i>	
On Performance Stability in LSM-based Storage Systems	449
<i>Chen Luo, Michael Carey</i>	
Hop-constrained s-t Simple Path Enumeration: Towards Bridging Theory and Practice	463
<i>You Peng, Ying Zhang, Xuemin Lin, Wenjie Zhang, Lu Qin, Jingren Zhou</i>	
Panorama: A Data System for Unbounded Vocabulary Querying over Video	477
<i>Yuhao Zhang, Arun Kumar</i>	
Planting Trees for scalable and efficient Canonical Hub Labeling.....	492
<i>Kartik Lakhoria, Rajgopal Kannan, Qing Dong, Viktor K Prasanna</i>	
Operationalizing Individual Fairness with Pairwise Fair Representations.....	506
<i>Preethi Lahoti, Krishna Gummadi, Gerhard Weikum</i>	
Optimizing Databases by Learning Hidden Parameters of Solid State Drives	519
<i>Aarati Kakraparthy, Jignesh Patel, Kwanghyun Park, Brian Kroth</i>	
BlazeIt: Optimizing Declarative Aggregation and Limit Queries for Neural Network-Based Video Analytics	533
<i>Daniel Kang, Peter D Bailis, Matei Zaharia</i>	
Joins on Samples: A Theoretical Guide for Practitioners	547
<i>Dawei Huang, Dong Young Yoon, Seth Pettie, Barzan Mozafari</i>	
Mining an "Anti-Knowledge Base" from Wikipedia Updates with Applications to Fact Checking and Beyond	561
<i>Georgios Karagiannis, Immanuel Trummer, Saehan Jo, Shubham Khandelwal, Xuezhi Wang, Cong Yu</i>	
Evaluating Persistent Memory Range Indexes	574
<i>Lucas Lersch, Xiangpeng Hao, Ismail Oukid, Tianzheng Wang, Thomas Willhalm</i>	

Research Papers

M.B.R.O.S.I.A: Providing Performant Virtual Resiliency for Distributed Applications	588
<i>Jonathan Goldstein, Ahmed S. Abdelhamid, Mike Barnett, Sebastian Burckhardt, Badrish Chandramouli, Darren Gehring, Niel I Lebeck, Christopher Meiklejohn, Umar Farooq Minhas, Ryan Newton, Rahee Peshawaria, Tal Zaccai, Irene Zhang</i>	
Efficient Shortest Path Index Maintenance on Dynamic Road Networks with Theoretical Guarantees	602
<i>Dian Ouyang, Long Yuan, Lu Qin, Lijun Chang, Ying Zhang, Xuemin Lin</i>	
ParPaRaw: Massively Parallel Parsing of Delimiter-Separated Raw Data	616
<i>Elias Stehle, Hans-Arno Jacobsen</i>	
Opportunities for Optimism in Contended Main-Memory Multicore Transactions.....	629
<i>Yihe Huang, William L Qian, Eddie Kohler, Barbara Liskov, Liuba Shrira</i>	
PM-LSH: A Fast and Accurate LSH Framework for High-Dimensional Approximate NN Search	643
<i>Bolong Zheng, Xi Zhao, Lianggui Weng, Hung Nguyen Quoc Viet, Hang Liu, Christian S Jensen</i>	
Hunting multiple bumps in graphs.....	656
<i>Yahui Sun, Jun Luo, Theodoros Lappas, Xiaokui Xiao, Bin Cui</i>	
Homogeneous Network Embedding for Massive Graphs via Reweighted Personalized PageRank .	670
<i>Renchi Yang, Jieming Shi, Xiaokui Xiao, Yin Yang, Sourav S Bhowmick</i>	
Pattern Functional Dependencies for Data Cleaning	684
<i>Abdulhakim Qahtan, Nan Tang, Mourad Ouzzani, Yang Cao, Michael Stonebraker</i>	
MEGA: Multi-View Semi-Supervised Clustering of Hypergraphs	698
<i>Joyce J Whang, Rundong Du, Sangwon Jung, Geon Lee, Barry Drake, Qingqing Liu, Seonggoo Kang, Haesun Park</i>	
MDedup: Duplicate Detection with Matching Dependencies.....	712
<i>Ioannis Koumarelas, Thorsten Papenbrock, Felix Naumann</i>	
Programmable View Update Strategies on Relations	726
<i>Van-Dang Tran, Hiroyuki Kato, Zhenjiang Hu</i>	
Amber: A Debuggable Dataflow System Based on the Actor Model	740
<i>Avinash Kumar, Zuozhi Wang, Shengquan Ni, Chen Li</i>	
Dynamic Speculative Optimizations for SQL Compilation in Apache Spark.....	754
<i>Filippo Schiavio, Daniele Bonetta, Walter Binder</i>	

Mind the Gap: An Experimental Evaluation of Imputation of Missing Values Techniques in Time Series	768
<i>Mourad Khayati, Alberto Lerner, Zakhar Tymchenko, Philippe Cudre-Mauroux</i>	

Research Papers

Graphite: A NUMA-aware HPC System for Graph Analytics Based on a new MPI*X Parallelism Model	783
<i>Mohammad Hasanzadeh Mofrad, Rami Melhem, Muhammad Yousuf Ahmad, Mohammad Hammoud</i>	
Personal Insights for Altering Decisions of Tree-based Ensembles over Time	798
<i>Nave Frost, Naama Boer, Daniel Deutch, Tova Milo</i>	
Answering Billion-Scale Label-Constrained Reachability Queries within Microsecond	812
<i>You Peng, Ying Zhang, Xuemin Lin, Lu Qin, Wenjie Zhang</i>	
Effective and Efficient Retrieval of Structured Entities	826
<i>Ruihong Huang, Shaoxu Song, Yunsu Lee, Jungho Park, Soo-Hyung Kim, Sungmin Yi</i>	
Micro-architectural Analysis of OLAP: Limitations and Opportunities.....	840
<i>Utku Sirin, Anastasia Ailamaki</i>	
Effective and Efficient Community Search over Large Heterogeneous Information Networks.....	854
<i>Yixiang Fang, Yixing Yang, Wenjie Zhang, Xuemin Lin, Xin Cao</i>	
ResilientDB: Global Scale Resilient Blockchain Fabric	868
<i>Suyash Gupta, Sajjad Rahnama, Jelle Hellings, Mohammad Sadoghi</i>	
Data-Parallel Query Processing on Non-Uniform Data	884
<i>Henning Funke, Jens Teubner</i>	
Evaluating Memory-Hard Proof-of-Work Algorithms on Three Processors	898
<i>Zonghao Feng, Qiong Luo</i>	
Approximate Summaries for Why and Why-not Provenance	912
<i>Seokki Lee, Bertram Ludaescher, Boris Glavic</i>	
PIDS: Attribute Decomposition for Improved Compression and Query Performance in Columnar Storage	925
<i>Hao Jiang, Chunwei Liu, Qi Jin, John Paparrizos, Aaron J Elmore</i>	
On Detecting Cherry-picked Trendlines.....	939
<i>Abolfazl Asudeh, H. V. Jagadish, You Wu, Cong Yu</i>	

Research Papers

Data-Driven Domain Discovery for Structured Datasets.....	953
<i>Masayo Ota*; Heiko Mueller; Juliana Freire ; Divesh Srivastava</i>	
Realtime Index-Free Single Source SimRank Processing on Web-Scale Graphs	966
<i>Jieming Shi*; Tianyuan Jin; Renchi Yang; Xiaokui Xiao; Yin Yang</i>	
Demand-Aware Route Planning for Shared Mobility Services	979
<i>Jiachuan Wang; Peng CHENG*; Libin Zheng; Chao Feng; Lei Chen; Xuemin Lin; Zheng Wang</i>	
DeepDB: Learn from Data, not from Queries!.....	992
<i>Benjamin Hilprecht*; Andreas Schmidt; Moritz Kulesa; Alejandro Molina; Kristian Kersting; Carsten Binnig</i>	
Data Migration using Datalog Program Synthesis.....	1006
<i>Yuepeng Wang*; Rushi Shah; Abby Criswell; Rong Pan; Isil Dillig</i>	
LiveGraph: A Transactional Graph Storage System with Purely Sequential Adjacency List Scans..	1020
<i>Xiaowei Zhu*; Guanyu Feng; Marco Serafini; Xiaosong Ma; Jiping Yu; Lei Xie; Ashraf Aboulnaga; Wenguang Chen</i>	
KBPearl: A Knowledge Base Population System Supported by Joint Entity and Relation Linking...	1035
<i>Xueling Lin*; Haoyang Li; Hao Xin; Zijian Li; Lei Chen</i>	
Compression of Uncertain Trajectories in Road Networks.....	1050
<i>Tianyi Li; Ruikai Huang; Lu Chen*; Christian S Jensen; Torben B Pedersen</i>	
Understanding and Benchmarking the Impact of GDPR on Database Systems	1064
<i>Supreeth Shastri*; Vinay Banakar; Melissa Wasserman; Arun Kumar; Vijay Chidambaram</i>	
LB+-Trees: Optimizing Persistent Index Performance on 3DXPoint Memory.....	1078
<i>Jihang Liu; Shimin Chen*; Lujun Wang</i>	
Enabling Low Tail Latency on Multicore Key-Value Stores.....	1091
<i>Lucas Lersch*; Ivan Schreter; Ismail Oukid; Wolfgang Lehner</i>	
Plato: Approximate Analytics over Compressed Time Series with Tight Deterministic Error Guarantees	1105
<i>Chunbin Lin*; Etienne Boursier; Yannis Papakonstantinou</i>	
Traversing Large Graphs on GPUs with Unified Memory	1119
<i>Prasun Gera*; Hyojong Kim; Piyush Sao; Hyesoon Kim; David Bader</i>	
Supporting Hard Queries over Probabilistic Preferences	1134
<i>Haoyue Ping*; Julia Stoyanovich; Benny Kimelfeld</i>	

Research Papers

Dash: Scalable Hashing on Persistent Memory.....	1147
<i>Baotong Lu, Xiangpeng Hao, Tianzheng Wang, Eric Lo</i>	
The PGM-index: a fully-dynamic compressed learned index with provable worst-case bounds	1162
<i>Paolo Ferragina, Giorgio Vinciguerra</i>	
Diagnosing Root Causes of Intermittent Slow Queries in Cloud Databases.....	1176
<i>Minghua Ma, Zheng Yin, Shenglin Zhang, Sheng Wang, Christopher Zheng, Xinhao Jiang, Hanwen Hu, Cheng Luo, Yilin Li, Nengjun Qiu, Feifei Li, Changcheng Chen, Dan Pei</i>	
Pangolin: An Efficient and Flexible Graph Mining System on CPU and GPU	1190
<i>Xuhao Chen, Roshan Dathathri, Gurbinder S Gill, Keshav Pingali</i>	
Quantifying TPC-H Choke Points and Their Optimizations.....	1206
<i>Markus Dreseler, Martin Boissier, Tilmann Rabl, Matthias Uflacker</i>	
Efficient Algorithms for Crowd-Aided Categorization.....	1221
<i>Yuanbing Li, Xian Wu, Yifei Jin, Jian Li, Guoliang Li</i>	
Set-valued Data Publication with Local Privacy: Tight Error Bounds and Efficient Mechanisms	1234
<i>Shaowei Wang, Yuqiu Qian, Jiachun Du, Wei Yang, Liusheng Huang, Hongli Xu</i>	
Translation of Array-Based Loops to Distributed Data-Parallel Programs.....	1248
<i>Leonidas Fegaras, Md Hasanuzzaman Noor</i>	
Incrementalization of Graph Partitioning Algorithms.....	1261
<i>Wenfei Fan, Muyang Liu, Chao Tian, Ruiqi Xu, Jingren Zhou</i>	
Optimizing Item and Subgroup Configurations for Social-Aware VR Shopping.....	1275
<i>Shao-Heng Ko, Hsu-Chao Lai, Hong-Han Shuai, Wang-Chien Lee, Philip Yu, De-Nian Yang</i>	
Efficient Confidentiality-Preserving Data Analytics over Symmetrically Encrypted Datasets	1290
<i>Savvas Savvides, Darshika Khandelwal, Patrick Eugster</i>	
Single Machine Graph Analytics on Massive Datasets Using Intel Optane DC Persistent Memory ..	1304
<i>Gurbinder S Gill, Roshan Dathathri, Loc Hoang, Ramesh Peri, Keshav Pingali</i>	

Research Papers

Atomic Commitment Across Blockchains	1319
<i>Victor Zakhary, Divy Agrawal, Amr El Abbadi</i>	
HydraList: A Scalable In-Memory Index Using Asynchronous Updates and Partial Replication	1332
<i>Ajit Mathew, Changwoo Min</i>	
eXtreme Modelling in Practice	1346
<i>A. Jesse Jiryu Davis, Max Hirschhorn, Judah Schvimer</i>	
Maximum Biclique Search at Billion Scale	1359
<i>Bingqing Lyu, Lu Qin, Xuemin Lin, Ying Zhang, Zhengping Qian, Jingren Zhou</i>	
ARDA: Automatic Relational Data Augmentation for Machine Learning	1373
<i>Nadiia Chepurko, Ryan Marcus, Emanuel Zraggen, Raul Castro Fernandez, Tim Kraska, David Karger</i>	
An LSM-based Tuple Compaction Framework for Apache AsterixDB	1388
<i>Wail Y Alkowiileet, Sattam Alsubaiee, Michael Carey</i>	
ADnEV: Cross-Domain Schema Matching using Deep Similarity Matrix Adjustment and Evaluation	1401
<i>Roe Shraga, Avigdor Gal, Haggai Roitman</i>	
Query Performance Prediction for Concurrent Queries using Graph Embedding	1416
<i>Xuanhe Zhou, Ji Sun, Guoliang Li, Jianhua Feng</i>	
Scalable, Near-Zero Loss Disaster Recovery for Distributed Data Stores	1429
<i>Ahmed Alquraan, Alex Kogan, Virendra Marathe, Samer Al-Kiswany</i>	
VHP: Approximate Nearest Neighbor Search via Virtual Hypersphere Partitioning.....	1443
<i>Kejing Lu, Hongya Wang, Wei Wang, Mineichi Kudo</i>	
IDAR: Fast Supergraph Search Using DAG Integration	1456
<i>Hyunjoon Kim, Seunghwan Min, Kunsoo Park, Xuemin Lin, Seok-Hee Hong, Wook-Shin Han</i>	
Guided Exploration of User Groups	1469
<i>Mariia Seleznova, Behrooz Omidvar-Tehrani, Sihem Amer-Yahia, Eric Simon</i>	
iDEC: Indexable Distance Estimating Codes for Approximate Nearest Neighbor Search	1483
<i>Long Gong, Huayi Wang, Mitsunori Ogihara, Jun Xu</i>	
Efficient Algorithms for Budgeted Influence Maximization on Massive Social Networks	1498
<i>Song Bian, Qintian Guo, Sibor Wang, Jeffrey Xu Yu</i>	
Mining Top-k Pairs of Correlated Subgraphs in a Large Network	1511

Arneish Prateek, Arijit Khan, Akshit Goyal, Sayan Ranu

FireLedger: A High Throughput Blockchain Consensus Protocol 1525
Yehonatan Buchnik, Roy Friedman

Put an Elephant into a Fridge: Optimizing Cache Efficiency for In-memory Key-value Stores..... 1540
Kefei Wang, Jian Liu, Feng Chen

Anytime Stochastic Routing with Hybrid Learning 1555
Simon Aagaard Pedersen, Bin Yang, Christian S Jensen

Understanding the Effect of Data Center Resource Disaggregation on Production DBMSs 1568
Qizhen Zhang, Yifan Cai, Xinyi Chen, Sebastian Angel, Ang Chen, Vincent Liu, Boon Thau Loo

Optimal Algorithms for Ranked Enumeration of Answers to Full Conjunctive Queries..... 1582
Nikolaos Tziavelis, Deepak Ajwani, Wolfgang Gatterbauer, Mirek Riedewald, Xiaofeng Yang

Sage: Parallel Semi-Asymmetric Graph Algorithms for NVRAMs 1598
Laxman Dhulipala, Charles McGuffey, Hongbo Kang, Yan Gu, Guy Blleloch, Phillip B Gibbons, Julian Shun

Research Papers

Pricing Influential Nodes in Online Social Networks.....	1614
<i>Yuqing Zhu, Jing Tang, Xueyan Tang</i>	
KClist++: A Simple Algorithm for Finding k-Clique Densest Subgraphs in Large Graphs.....	1628
<i>Bintao Sun, Maximilien Dansich, T-H. Hubert Chan, Mauro Sozio</i>	
Dynamic Interleaving of Content and Structure for Robust Indexing of Semi-Structured Hierarchical Data	1641
<i>Kevin Wellenzohn, Michael H Böhlen, Sven Helmer</i>	
ChiSeL: Graph Similarity Search using Chi-Squared Statistics in Large Probabilistic Graphs	1654
<i>Shubhangi Agarwal, Sourav Dutta, Arnab Bhattacharya</i>	
Fast Incremental Discovery of Pointwise Order Dependencies	1669
<i>Zijing Tan, Ai Ran, Shuai Ma, Sheng Qin</i>	
Approximate Denial Constraints	1682
<i>Ester Livshits, Alireza Heidari, Ihab F Ilyas, Benny Kimelfeld</i>	
Sharing Opportunities for OLTP Workloads in Different Isolation Levels.....	1696
<i>Robin Rehrmann, Carsten Binnig, Alexander Boehm, Kihong Kim, Wolfgang Lehner</i>	
Biclustering and Boolean Matrix Factorization in Data Streams	1709
<i>Stefan Neumann, Pauli Miettinen</i>	
Effective and Efficient Relational Community Detection and Search in Large Dynamic Heterogeneous Information Networks	1723
<i>Xun Jian, Yue Wang, Lei Chen</i>	
Natural language to SQL: Where are we today?	1737
<i>Hyeonji Kim, Byeong-Hoon So, Wook-Shin Han, Hongrae Lee</i>	
Accelerating Truss Decomposition on Heterogeneous Processors.....	1751
<i>Yulin Che, Zhuohang Lai, Shixuan Sun, Yue Wang, Qiong Luo</i>	
Searching a Database of Source Codes Using Contextualized Code Search	1765
<i>Rohan Mukherjee, Chris Jermaine, Swarat Chaudhuri</i>	
Data Stream Event Prediction Based on Timing Knowledge and State Transitions	1779
<i>Yan Li, Tingjian Ge, Cindy Chen</i>	
Shared Arrangements: practical inter-query sharing for streaming dataflows.....	1793
<i>Frank McSherry, Andrea Lattuada, Malte Schwarzkopf, Timothy Roscoe</i>	

Research Papers

SmartBench: A Benchmark For Data Management In Smart Spaces.....	1807
<i>Peeyush Gupta, Michael Carey, Sharad Mehrotra, Roberto Yus</i>	
Series2Graph: Graph-based Subsequence Anomaly Detection for Time Series	1821
<i>Paul Boniol, Themis Palpana</i>	
Sato: Contextual Semantic Type Detection in Tables.....	1835
<i>Dan Zhang, Yoshihiko Suhara, Jinfeng Li, Madelon Hulsebos, Çağatay Demiralp, Wang-Chiew Tan</i>	
TransNet: Training Privacy-Preserving Neural Network over Transformed Layer	1849
<i>Qijian He, Wei Yang, Bingren Chen, Yangyang Geng, Liusheng Huang</i>	
Capturing Associations in Graphs.....	1863
<i>Wenfei Fan, Ruochun Jin, Muyang Liu, Ping Lu, Chao Tian, Jingren Zhou</i>	
Dynamic Parameter Allocation in Parameter Servers	1877
<i>Alexander Renz-Wieland, Rainer Gemulla, Steffen Zeuch, Volker Markl</i>	
Adopting Worst-Case Optimal Joins in Relational Database Systems	1891
<i>Michael Freitag, Maximilian Bandle, Tobias Schmidt, Alfons Kemper, Thomas Neumann</i>	
A workload-adaptive mechanism for linear queries under local differential privacy	1905
<i>Ryan McKenna, Raj Kumar Maity, Arya Mazumdar, Gerome Miklau</i>	
SPORES: Sum-Product Optimization via Relational Equality Saturation for Large Scale Linear Algebra	1919
<i>Yisu R Wang, Shana Hutchison, Dan Suciu, Bill G Howe, Jonathan L Leang</i>	
Data Market Platforms: Trading Data Assets to Solve Data Problems	1933
<i>Raul Castro Fernandez, Pranav Subramaniam, Michael Franklin</i>	
Baran: Effective Error Correction via a Unified Context Representation and Transfer Learning	1948
<i>Mohammad Mahdavi, Ziawasch Abedjan</i>	
Relational Data Synthesis using Generative Adversarial Networks: A Design Space Exploration....	1962
<i>Ju Fan, Tongyu Liu, Guoliang Li, Junyou Chen, Yuwei Shen, Xiaoyong Du</i>	
Leaper: A Learned Prefetcher for Cache Invalidation in LSM-tree based Storage Engines	1976
<i>Lei Yang, HONG WU, Tieying Zhang, Xuntao Cheng, Feifei Li, Lei Zou, Yujie Wang, Rongyao Chen, Jianying Wang, Gui Huang</i>	

Approximate Selection with Guarantees using Proxies	1990
<i>Daniel Kang, Edward Gan, Peter D Bailis, Tatsunori Hashimoto, Matei Zaharia</i>	
2R: Efficiently Isolating Cold Pages in Flash Storages.....	2004
<i>Minji Kang, Soyee Choi, Gihwan Oh, Sang Won Lee</i>	
Knowledge Translation.....	2018
<i>Bahar Ghadiri Bashardoost, Renée J. Miller, Kelly Lyons, Fatemeh Nargesian</i>	
Towards Scalable Dataframe Systems.....	2033
<i>Devin Petersohn, Stephen Macke, Doris Xin, William Ma, Doris Lee, Xiangxi Mo, Joseph Gonzalez, Joseph M Hellerstein, Anthony Joseph, Aditya Parameswaran</i>	
Aria: A Fast and Practical Deterministic OLTP Database	2047
<i>Yi Lu, Xiangyao Yu, Lei Cao, Samuel Madden</i>	
The Computation of Optimal Subset Repairs	2061
<i>Dongjing Miao, Zhipeng Cai, Jianzhong Li, Xiangyu Gao, Xianmin Liu</i>	
Pytheas: Pattern-based Table Discovery in CSV Files	2075
<i>Christina Christodoulakis, Eric B Munson, Moshe Gabel, Angela Demke Brown, Renée J. Miller</i>	
Privacy Preserving Vertical Federated Learning for Tree-based Models.....	2090
<i>Yuncheng Wu, Shaofeng Cai, Xiaokui Xiao, Gang Chen, Beng Chin Ooi</i>	
Topic-based Community Search over Spatial-Social Networks.....	2104
<i>Ahmed Al-Baghdadi, Xiang Lian</i>	
LOG-Means: Efficiently Estimating the Number of Clusters in Large Datasets.....	2118
<i>Manuel Fritz, Michael Behringer, Holger Schwarz</i>	
Efficient Oblivious Database Joins.....	2132
<i>Simeon Krastnikov, Florian Kerschbaum, Douglas Stebila</i>	
Evaluating Top-k Queries with Inconsistency Degrees.....	2146
<i>Ousmane Issa, Angela Bonifati, Farouk Toumani</i>	
Cerebro: A Data System for Optimized Deep Learning Model Selection	2159
<i>Supun Nakandala, Yuhao Zhang, Arun Kumar</i>	
CoopStore: Optimizing Precomputed Summaries for Aggregation	2174
<i>Edward Gan, Peter D Bailis, Moses Charikar</i>	
Fast Subtrajectory Similarity Search in Road Networks under Weighted Edit Distance Constraints	2188
<i>Satoshi Koide, Chuan Xiao, Yoshiharu Ishikawa</i>	
SimTab: Accuracy-Guaranteed SimRank Queries through Tighter Confidence Bounds and Multi-Armed Bandits	2202
<i>Yu Liu, Lei Zou, Qian Ge, Zhewei Wei</i>	
Efficiently Approximating Selectivity Functions using Low Overhead Regression Models	2215
<i>Anshuman Dutt, Chi Wang, Vivek Narasayya, Surajit Chaudhuri</i>	

Identifying Insufficient Data Coverage in Databases with Multiple Relations	2229
<i>Yin Lin, Yifan Guan, Abolfazl Asudeh, H. V. Jagadish</i>	
Continuously Monitoring Alternative Shortest Paths on Road Networks	2243
<i>Lingxiao Li, Muhammad Aamir Cheema, Mohammed Eunus Ali, Hua Lu, David Taniar</i>	
Hypergraph Motifs: Concepts, Algorithms, and Discoveries.....	2256
<i>Geon Lee, Jihoon Ko, Kijung Shin</i>	
Hitting Set Enumeration with Partial Information for Unique Column Combination Discovery	2270
<i>Johann Birnick, Thomas Bläsius, Tobias Friedrich, Felix Naumann, Thorsten Papenbrock, Martin Schirneck</i>	
SSTD: A Distributed System on Streaming Spatio-Textual Data	2284
<i>Yue Chen, Zhida Chen, Gao Cong, Ahmed Mahmood, Walid Aref</i>	
Continuous Prefetch for Interactive Data Applications	2297
<i>Haneen Mohammed, Ziyun Wei, Eugene Wu, Ravi Netravali</i>	
Efficient and Effective Similar Subtrajectory Search with Deep Reinforcement Learning	2312
<i>Zheng WANG, Cheng Long, Gao Cong, Yiding Liu</i>	
A Benchmarking Study of Embedding-based Entity Alignment for Knowledge Graphs.....	2326
<i>Zequn Sun, Qingheng Zhang, Wei Hu, Chengming Wang, Muhao Chen, Farahnaz Akrami, Chengkai Li</i>	
Effectively Learning Spatial Indices	2341
<i>Jianzhong Qi, Guanli Liu, Christian S Jensen, Lars Kulik</i>	
Stable Learned Bloom Filters for Data Streams	2355
<i>Qiyu LIU, Libin Zheng, Yanyan Shen, Lei Chen</i>	
Auto-Transform: Learning-to-Transform by Patterns	2368
<i>Zhongjun Jin, Yeye He, Surajit Chaudhuri</i>	
Magic mirror in my hand, which is the best in the land? An Experimental Evaluation of Index Selection Algorithms	2382
<i>Jan Kossmann, Stefan Halfpap, Marcel Jankrift, Rainer Schlosser</i>	
MorphStore: Analytical Query Engine with a Holistic Compression-Enabled Processing Model	2396
<i>Patrick Damme, Annett Ungethüm, Johannes Pietrzyk, Alexander Krause, Dirk Habich, Wolfgang Lehner</i>	
Fast and Effective Distribution-Key Recommendation for Amazon Redshift	2411
<i>Panos Parchas, Yonatan Naamad, Peter Van Bouwel, Christos Faloutsos, Michalis Petropoulos</i>	
Sieve: A Middleware Approach to Scalable Access Control for Database Management Systems....	2424
<i>Primal Pappachan, Roberto Yus, Sharad Mehrotra, Johann-Christoph Freytag</i>	
Cloudburst: Stateful Functions-as-a-Service	2438
<i>Vikram Sreekanti, Chenggang Wu, Charles Lin, Johann Schleier-Smith, Joseph Gonzalez, Joseph M Hellerstein, Alexey Tumanov</i>	

ODIN: Automated Drift Detection and Recovery in Video Analytics	2453
<i>Abhijit Suprem, Joy Arulraj, Calton Pu, Joao E Ferreira</i>	
Maximizing the Reduction Ability for Near-maximum Independent Set Computation	2466
<i>Chengzhi Piao, Weiguo Zheng, Yu Rong, Hong Cheng</i>	
Fair Task Assignment in Spatial Crowdsourcing	2479
<i>Zhao CHEN, Peng CHENG, Lei Chen, Xuemin Lin, Cyrus Shahabi</i>	
Distributed Subgraph Counting: A General Approach	2493
<i>Hao Zhang, Jeffrey Xu Yu, Yikai Zhang, Kangfei Zhao, Hong Cheng</i>	
Scrutinizer: A Mixed-Initiative Approach to Large-Scale, Data-Driven Claim Verification	2508
<i>Georgios Karagiannis, Mohammed Saeed, Paolo Papotti, Immanuel Trummer</i>	
Detecting and Preventing Confused Labels in Crowdsourced Data	2522
<i>Evgeny Krivosheev, Siarhei Bykau, Fabio Casati, Sunil Prabhakar</i>	
Ordering Heuristics for k-clique Listing	2536
<i>Ronghua Li, Sen Gao, Lu Qin, Guoren Wang, Weihua Yang, Jeffrey Xu Yu</i>	
Deep or Simple Models for Semantic Tagging? It Depends on your Data	2549
<i>Jinfeng Li, Yuliang Li, Xiaolan Wang, Wang-Chiew Tan</i>	
Do the Best Cloud Configurations Grow on Trees? An Experimental Evaluation of Black Box Algorithms for Optimizing Cloud Workloads.....	2563
<i>Muhammad Bilal, Marco Serafini, Marco Canini, Rodrigo Rodrigues</i>	
Finding Large Diverse Communities on Networks: The Edge Maximum k^* -Partite Clique	2576
<i>Alexander Zhou, Yue Wang, Lei Chen</i>	
Practical Client-side Replication: Weak Consistency Semantics for Insecure Settings	2590
<i>Albert van der Linde, João Leitão, Nuno Preguiça</i>	
Approximate Partition Selection for Big-Data Workloads using Summary Statistics	2606
<i>Kexin Rong, Yao Lu, Peter Bailis, Srikanth Kandula, Philip Levis</i>	
Meet Me Halfway: Split Maintenance of Continuous Views.....	2620
<i>Christian Winter, Tobias Schmidt, Thomas Neumann, Alfons Kemper</i>	
μ Tree: a Persistent B+-Tree with Low Tail Latency	2634
<i>Youmin Chen, Youyou Lu, Kedong Fang, Qing Wang, Jiwu Shu</i>	
FSST: Fast Random Access String Compression.....	2649
<i>Peter Boncz, Thomas Neumann, Viktor Leis</i>	
Mosaic: A Budget-Conscious Storage Engine for Relational Database Systems.....	2662
<i>Lukas Vogel, Alexander van Renen, Satoshi Imamura, Viktor Leis, Thomas Neumann, Alfons Kemper</i>	
Many-Core Clique Enumeration with Fast Set Intersections	2676
<i>Jovan Blanuša, Radu Stoica, Paolo Ienne, Kubilay Atasu</i>	
SAQE: Practical Privacy-Preserving Approximate Query Processing for Data Federations	2691

Johes Bater, Yongjoo Park, Xi He, Xiao Wang, Jennie Rogers

Rank Aggregation Algorithms for Fair Consensus	2706
<i>Caitlin Kuhlman, Elke Rundensteiner</i>	
Sentinel: Universal Analysis and Insight for Data Systems	2720
<i>Brad Glasbergen, Michael Abebe, Khuzaima Daudjee, Amit Levi</i>	
Optimizing DNN Computation Graph using Graph Substitutions	2734
<i>Jingzhi Fang, Yanyan Shen, Yue Wang, Lei Chen</i>	
ATHENA++: Natural Language Querying for Complex Nested SQL Queries	2747
<i>Jaydeep Sen, Chuan Lei, Abdul Quamar, Fatma Ozcan, Vasilis Efthymiou, Ayushi Dalmia, Greg Stager, Ashish Mittal, Diptikalyan Saha, Karthik Sankaranarayanan</i>	
Collecting and Analyzing Data Jointly from Multiple Services under Local Differential Privacy	2760
<i>Min Xu, Bolin Ding, Tianhao Wang, Jingren Zhou</i>	
IsoDiff: Debugging Anomalies Caused by Weak Isolation	2773
<i>Yifan Gan, Xueyuan Ren, Drew Ripberger, Spyros Blanas, Yang Wang</i>	
Suffix Rank: a new scalable algorithm for indexing large string collections	2787
<i>Marina Barsky, Jonathan Gabor, Mariano Consens, Alex Thomo</i>	

Demonstrations

Demand-based Sensor Data Gathering with Multi-Query Optimization	2801
<i>Julius Hülsmann, Jonas Traub, Volker Markl</i>	
CheetahVIS: A Visual Analytical System for Large Urban Bus Data.....	2805
<i>Wentao Ning, Qiandong Tang, Yi Zhao, Chuan Yang, Xiaofeng Wang, Teng Wang, Haotian Liu, Chaozu Zhang, Zhiyuan Zhou, Qiaomu Shen, Bo Tang</i>	
UNMASQUE: A Hidden SQL Query Extractor	2809
<i>Kapil Khurana, Jayant Haritsa</i>	
G3: When Graph Neural Networks Meet Parallel Graph Processing Systems on GPUs	2813
<i>Husong Liu, Shengliang Lu, Xinyu Chen, Bingsheng He</i>	
PiBench Online: Interactive Benchmarking of Persistent Memory Indexes	2817
<i>Xiangpeng Hao, Lucas Lersch, Tianzheng Wang, Ismail Oukid</i>	
VisClean: Interactive Cleaning for Progressive Visualization	2821
<i>Yuyu Luo, Chengliang Chai, Xuedi Qin, Nan Tang, Guoliang Li</i>	
IMO: A Toolbox for Simulating and Querying "Infected" Moving Objects	2825
<i>Jianqiu Xu, Hua Lu, Zhifeng Bao</i>	
COUNTATA: Dataset Labeling Using Pattern Counts.....	2829
<i>Yuval Moskovitch, H. Jagadish</i>	
A Demonstration of Willump: A Statistically-Aware End-to-end Optimizer for Machine Learning ...	2833
<i>Inference</i>	
<i>Peter Kraft, Daniel Kang, Deepak Narayanan, Shoumik Palkar, Peter Bailis, Matei Zaharia</i>	
Ease.ml/snoopy in Action: Towards Automatic Feasibility Analysis for Machine Learning Application Development.....	2837
<i>Cedric Renggli, Luka Rimanic, Luka Kolar, Wentao Wu, Ce Zhang</i>	
DeepTrack: Monitoring and Exploring Spatio-Temporal Data – A Case of Tracking COVID-19	2841
<i>Yuyu Luo, Wenbo Li, Tianyu Zhao, Xiang Yu, Lixi Zhang, Guoliang Li, Nan Tang</i>	
DF-Toolkit: Interacting with Low-Level Database Storage	2845
<i>James Wagner, Alexander Rasin, Karen Heart, Tanu Malik, Jonathan Grier</i>	

Like Water and Oil: With a Proper Emulsifier, Query Compilation and Data Parallelism Will Mix Well	2849
<i>Henning Funke, Jens Teubner</i>	
SQUARES : A SQL Synthesizer Using Query Reverse Engineering.....	2853
<i>Pedro Orvalho, Miguel Terra-Neves, Miguel Ventura, Ruben Martins, Vasco Manquinho</i>	
SuccinctEdge: A Succinct RDF Store for Edge Computing	2857
<i>Weiqin Xu, Olivier CURE, Philippe Calvez</i>	
SuDocu: Summarizing Documents by Example	2861
<i>Anna Fariha, Matteo Brucato, Peter Haas, Alexandra Meliou</i>	
CONCIERGE: Improving Constrained Search Results by Data Melioration	2865
<i>Ido Guy, Tova Milo, Slava Novgorodov, Brit Youngmann</i>	
Demonstrating the Voice-Based Exploration of Large Data Sets with CiceroDB-Zero.....	2869
<i>Immanuel Trummer</i>	
FASTS: A Satisfaction-Boosting Bus Scheduling Assistant	2873
<i>Songsong Mo, Zhifeng Bao, Baihua Zheng, Zhiyong Peng</i>	
Vaas: Video Analytics At Scale	2877
<i>Favyen Bastani, Oscar Moll, Samuel Madden</i>	
sPaQLTools: A Stochastic Package Query Interface for Scalable Constrained Optimization.....	2881
<i>Matteo Brucato, Miro Mannino, Azza Abouzied, Peter Haas, Alexandra Meliou</i>	
ActiveDeeper: A Model-based Active Data Enrichment System.....	2885
<i>Liang Zhao, Qingcan Li, Pei Wang, Jiannan Wang, Eugene Wu</i>	
RDFFrames: Knowledge Graph Access for Machine Learning Tools	2889
<i>Aisha Mohamed, Ghadeer Abuoda, Abdurrahman Ghanem, Zoi Kaoudi, Ashraf Aboulnaga</i>	
Scalable, Resilient and Configurable Permissioned Blockchain Fabric.....	2893
<i>Sajjad Rahnama, Suyash Gupta, Thamir Qadah, Jelle Hellings, Mohammad Sadoghi</i>	
BIRDS: Programming view update strategies in Datalog	2897
<i>Van-Dang Tran, Hiroyuki Kato, Zhenjiang Hu</i>	
Apache IoTDB: Time-series database for Internet of Things.....	2901
<i>Chen Wang, Huang Xiangdong, Jialin Qiao, Jianmin Wang, Jiaguang Sun, Kevin McGrail, Julian Feinauer, Jinrui Zhang, Peng Wang, Jinrui Zhang, Rong Kang, Tian Jiang, Lei Rui, Jun Yuan</i>	
Evaluating Ridesharing Algorithms using the Jargo Real-Time Stochastic Simulator.....	2905
<i>James Pan, Guoliang Li, Yong Wang</i>	
BitFun: Fast Answers to Queries with Tunable Functions in Geospatial Array DBMS	2909
<i>Ramon Antonio Rodrigues Zalipynis</i>	
SPHINX: A System for Metapath-based Entity Exploration in Heterogeneous Information Networks	2913

<i>Serafeim Chatzopoulos, Kostas Patroumpas, Alexandros Zeakis, Thanasis Vergoulis, Dimitrios Skoutas</i> ExplainED: Explanations for EDA Notebooks	2917
<i>Daniel Deutch, Amir Gilad, Tova Milo, Amit Somech</i>	
MuSe: Multiple Deletion Semantics for Data Repair	2921
<i>Amir Gilad, Yihao Hu, Daniel Deutch, Sudeepa Roy</i>	
Tabula in Action: A Sampling Middleware for Interactive Geospatial Visualization Dashboards	2925
<i>Jia Yu, Kanchan Chowdhury, Mohamed Sarwat</i>	
X ² R ² : a Tool for Explainable and Explorative Reidentification Risk Analysis	2929
<i>Tom Rolandus Hagedoorn, Rohit Kumar, Francesco Bonchi</i>	
Obi-Wan: Ontology-Based RDF Integration of Heterogeneous Data	2933
<i>Maxime Buron, François Goasdoué, Ioana Manolescu, Marie-Laure Mugnier</i>	
CrocodileDB in Action: Resource-Efficient Query Execution by Exploiting Time Slackness	2937
<i>Dixin Tang, Zechao Shang, Aaron Elmore, Sanjay Krishnan, Michael Franklin</i>	
GraphAn: Graph-based Subsequence Anomaly Detection	2941
<i>Paul Boniol, Themis Palpanas, Mohammed Meftah, Emmanuel Remy</i>	
LMFAO: An Engine for Batches of Group-By Aggregates.....	2945
<i>Maximilian Schleich, Dan Olteanu</i>	
ESTOCADA: Towards Scalable Polystore Systems	2949
<i>Rana Alotaibi, Bogdan Cautis, Alin Deutsch, Moustafa Latrache, Ioana Manolescu, Yifei Yang</i>	
Demonstration of Interactive Runtime Debugging of Distributed Dataflows in Texera	2953
<i>Zuozhi Wang, Avinash Kumar, Shengquan Ni, Chen Li</i>	
DeepTRANS: A Deep Learning System for Public Bus Travel Time Estimation using Traffic Forecasting	2957
<i>Luan Tran, Min Young Mun, Matthew Lim, Jonah Yamato, Nathan Huh, Cyrus Shahabi</i>	
Demonstration of ScroogeDB: Getting More Bang For the Buck with Deterministic Approximation in the Cloud	2961
<i>Saehan Jo, Jialing Pei, Immanuel Trummer</i>	
Scrutinizer: Fact Checking Statistical Claims	2965
<i>Georgios Karagiannis, Mohammed Saeed, Paolo Papotti, Immanuel Trummer</i>	
SciLens News Platform: A System for Real-Time Evaluation of News Articles	2969
<i>Angelika Romanou, Panayiotis Smeros, Carlos Castillo, Karl Aberer</i>	
HDAG-Explorer: A System for Hierarchical DAG Summarization and Exploration	2973
<i>Xuliang Zhu, Xin Huang, Jinbin Huang, Byron Choi, Jianliang Xu</i>	
Orca-SR: A Real-Time Traffic Engineering Framework leveraging Similarity Joins.....	2977
<i>Jees Augustine, Suraj Shetiya, Abolfazl Asudeh, Saravanan Thirumuruganathan, Azade Nazi, Nan Zhang, Gautam Das, Divesh Srivastava</i>	

nKV in Action: Accelerating KV-Stores on NativeComputational Storage with Near-Data Processing	2981
<i>Tobias Vincon, Lukas Weber, Arthur Bernhard, Andreas Koch, Iliia Petrov, Christian Knoedler, Sergey Hardock, Sajjad Tamimi, Christian Riegger</i>	
Demonstration of Inferring Causality from Relational Databases with CaRL	2985
<i>Moe Kayali, Babak Salimi, Dan Suciu</i>	
SQL for Data Scientists: Designing SQL Tutorials for Scalable Online Teaching	2989
<i>Uwe Roehm, Lexi Brent, Tim Dawborn, Bryn Jeffries</i>	
Debugging Large-Scale Data Science Pipelines using Dagger.....	2993
<i>El Kindi Rezig, Ashrita Brahmaroutu, Nesime Tatbul, Mourad Ouzzani, Nan Tang, Timothy Mattson, Samuel Madden, Michael Stonebraker</i>	
I-Rex: An Interactive Relational Query Explainer for SQL	2997
<i>Zhengjie Miao, Tiangang Chen, Alexander Bendeck, Kevin Day, Sudeepa Roy, Jun Yang</i>	
PANDA: Policy-aware Location Privacy for Epidemic Surveillance	3001
<i>Yang Cao, Shun Takagi, Yonghui Xiao, Li Xiong, Masatoshi Yoshikawa</i>	

Industrial and Applications

PyTorch Distributed: Experiences on Accelerating Data Parallel Training	3005
<i>Shen Li, Yanli Zhao, Rohan Varma, Omkar Salpekar, Pieter Noordhuis, Teng Li, Adam Paszke, Jeff Smith, Brian Vaughan, Pritam Damania, Soumith Chintala</i>	
Towards Multi-way Join Aware Optimizer in SAP HANA	3019
<i>Sungheun Wi, Wook-Shin Han, Chuho Chang, Kihong Kim</i>	
InvaliDB: Scalable Push-Based Real-Time Queries on Top of Pull-Based Databases	3032
<i>Wolfram Wingerath, Felix Gessert, Norbert Ritter</i>	
Automated Generation of Materialized Views in Oracle.....	3046
<i>Rafi Ahmed, Randall Bello, Andrew Witkowski, Praveen Kumar</i>	
Native JSON Datatype Support: Maturing SQL and NoSQL convergence in Oracle Database.....	3059
<i>Zhen Hua Liu, Beda Hammerschmidt, Douglas McMahon, Hui Chang, Ying Lu, Josh Spiegel, Alfonso Colunga Sosa, srikrishnan suresh, Geeta Arora, vikas Arora</i>	
TiDB: A Raft-based HTAP Database	3072
<i>Dongxu Huang, Qi Liu, Qiu Cui, Zhuhe Fang, Xiaoyu Ma, Fei Xu, Li Shen, Liu Tang, Yuxing Zhou, Menglong Huang, Wan Wei, Cong Liu, Jian Zhang, Jianjun Li, Xuelian Wu, Lingyu Song, Ruoxi Sun, Shuaipeng Yu, Lei Zhao, Nicholas Cameron, Liqun Pei, Xin Tang</i>	
A system design for elastically scaling transaction processing engines in virtualized servers	3085
<i>Angelos Christos Anadiotis, Raja Appuswamy, Anastasia Ailamaki, Ilan Bronshtein, Hillel Avni, David Dominguez-Sal, Shay Goikhman, Eliezer Levy</i>	

Industrial Strength OLTP Using Main Memory and Many Cores.....	3099
<i>Hillel Avni, Alisher Aliev, Oren Amor, Aharon Avitzur, Ilan Bronshtein, Eli Ginot, Shay Goikhman, Eliezer Levy, Idan Levy, Fuyang Lu, Liran Mishali, Yeqin Mo, Nir Pachter, Dima Sivov, Vinoth Veeraraghavan, Vladi Vexler, Lei Wang, Peng Wang</i>	
Asymmetric-Partition Replication for Highly Scalable Distributed Transaction Processing in Practice	3112
<i>Juchang Lee, Hyejeong Lee, Seongyun Ko, Kyu Hwan Kim, Mihnea ANDREI, Friedrich Keller, Wook-Shin Han</i>	
AGL: A Scalable System for Industrial-purpose Graph Machine Learning	3125
<i>Dalong Zhang, Xin Huang, ziqi liu, Jun Zhou, Zhiyang Hu, Xianzheng Song, Zhibang Ge, Lin Wang, Zhiqiang Zhang, Yuan Qi</i>	
LedgerDB: A Centralized Ledger Database for Universal Audit and Verification	3138
<i>Xinying Yang, Yuan Zhang, Sheng Wang, Benquan Yu, Feifei Li, Yize Li, Wenyuan Yan</i>	
AnalyticDB-V: A Hybrid Analytical Engine Towards Query Fusion for Structured and Unstructured Data	3152
<i>Chuangxian Wei, Bin Wu, Sheng Wang, Renjie Lou, Chaoqun Zhan, Feifei Li, Yuanzhe Cai</i>	
Oracle AutoML: A Fast and Predictive AutoML Pipeline	3166
<i>Anatoly Yakovlev, Hesam Fathi Moghadam, Ali Moharrer, Jingxiao Cai, Nikan Chavoshi, Venkatanathan Varadarajan, Sandeep Agrawal, Tomas Karnagel, Sam Idicula, Sanjay Jinturkar, Nipun Agarwal)</i>	
Monarch: Google's Planet-Scale In-Memory Time Series Database.....	3181
<i>Colin Adams, Luis Alonso, Benjamin Atkin, John Banning, Sumeer Bhola, Rick Buskens, Ming Chen, Xi Chen, Yoo Chung, Qin Jia, Nick Sakharov, George Talbot, Nick Taylor, Adam Tart</i>	
Concurrent Updates to Pages with Fixed-Size rows Using Lock-Free Algorithms.....	3195
<i>Raghavendra Thallam Kodandaramaih, Hanuma Kodavalla, Girish Venkataramanappa</i>	
POLARIS: The Distributed SQL Engine in Azure Synapse	3204
<i>Josep Aguilar Saborit, Raghu Ramakrishnan</i>	
MyRocks: LSM-Tree Database Storage Engine Serving Facebook's Social Graph	3217
<i>Yoshinori Matsunobu, Siying Dong, Herman Lee</i>	
Helios: Hyperscale Indexing for the Cloud & Edge	3231
<i>Rahul Potharaju, Terry Kim, Wentao Wu, Vidip Acharya, Steve Suh, Andrew Fogarty, Apoorve Dave, Sinduja Ramanujam, Tomas Talius, Lev Novik, Raghu Ramakrishnan</i>	
Replication at the Speed of Change – a Fast, Scalable Replication Solution for Near Real-Time HTAP Processing	3245
<i>Dennis Butterstein, Daniel Martin, Knut Stolze, Felix Beier, Jia Zhong, Lingyun Wang</i>	
Exploiting Domain Knowledge to address Multi-Class Imbalance and a Heterogeneous Feature Space in Classification Tasks for Manufacturing Data	3258
<i>Vitali Hirsch, Peter Reimann, Bernhard Mitschang</i>	
Alibaba Hologres: A Cloud-Native Service for Hybrid Serving/Analytical Processing	3272
<i>Xiaowei Jiang, Yuejun Hu, Yu Xiang, Guangran Jiang, Xiaojun Jin, Chen Xia, Weihua Jiang, Jun Yu, Haitao Wang, Yuan Jiang, Jihong Ma, Li Su, Kai Zeng</i>	

DIAMetrics: Benchmarking Query Engines at Scale	3285
<i>Anja Gruenheid, Shaleen Deep, Kruthi Nagaraj, Hiro Naito, Jeff Naughton, Stratis Viglas</i>	
Db2 Event Store: A Purpose-Built IoT Database Engine	3299
<i>Christian Garcia-Arellano, Adam Storm, David Kalmuk, Hamdi Roumani, Ronald Barber, Yuanyuan Tian, Richard Sidle, Fatma Ozcan, Matthew Spilchen, Josh Tiefenbach, Daniel Zilio, Lan Pham, Kostas Rakopoulos, Alexander Cheung, Darren Pepper, Imran Sayyid, Gidon Gershinsky, Gal Lushi, Hamid Pirahesh</i>	
F1 Lightning: HTAP as a Service.....	3313
<i>Jiacheng Yang, Ian Rae, Jun Xu, Jeff Shute, Zhan Yuan, Kelvin Lau, Qiang Zeng, Xi Zhao, Jun Ma, Ziyang Chen, Yuan Gao, Qilin Dong, Junxiong Zhou, Jeremy Wood, Goetz Graefe, Jeff Naughton, John Cieslewicz</i>	
AutoToken: Predicting Peak Parallelism for Big Data Analytics at Microsoft.....	3326
<i>Rathijit Sen, Alekh Jindal, Hiren Patel, Shi Qiao</i>	
A Drop-in Middleware for Serializable DB Clustering across Geo-distributed Sites.....	3340
<i>Enrique Saurez, Bharath Balasubramanian, Richard Schlichting, Brendan Tschaen, Shankaranarayanan Puzhavakath Narayanan, Zhe Huang, Umakishore Ramachandran</i>	
Improving Reproducibility of Data Science Pipelines through Transparent Provenance Capture ...	3354
<i>Lukas Rupprecht, James Davis, Constantine Arnold, Yaniv Gur, Deepavali Bhagwat</i>	
Conversational BI: An Ontology-Driven ConversationSystem for Business Intelligence Applications	3369
<i>Abdul Quamar, Fatma Ozcan, Dorian Miller, Robert Moore, Rebecca Niehus, Jeffrey Kreulen</i>	
Magnet: Push-based Shuffle Service for Large-scale Data Processing.....	3382
<i>Min Shen, Ye Zhou, Chandni Singh</i>	
Leveraging Organizational Resources to Adapt Models to New Data Modalities	3396
<i>Sahaana Suri, Abishek Sethi, Girija Narlikar, Neslihan Bulut, Raghuvveer Chanda, Sugato Basu, Pradyumna Narayana, Peter Bailis, Christopher Re, Yemao Zeng</i>	
Delta Lake: High-Performance ACID Table Storage over Cloud Object Stores	3411
<i>Michael Armbrust, Tathagata Das, Sameer Paranjpye, Reynold Xin, Shixiong Zhu, Ali Ghodsi, Burak Yavuz, Mukul Murthy, Joseph Torres, Liwen Sun, Peter Boncz, Mostafa Mokhtar, Herman van Hovell, Adrian Ionescu, Alicja Luszczak, Michal Switakowski, Takuya Ueshin, Xiao Li, Michal Szafranski, Pieter Senster, Matei Zaharia</i>	

Tutorials

Robust Query Processing: Mission Possible	3425
<i>Jayant Haritsa</i>	
Data Collection and Quality Challenges for Deep Learning.....	3429
<i>Steven Whang, Jae-Gil Lee</i>	
Table Extraction and Understanding for Scientific and Enterprise Applications	3433
<i>Doug Burdick, Marina Danilevsky, Alexandre Evfimievski, Yannis Katsis, Nancy X.R. Wang</i>	

Similarity Query Processing for High-Dimensional Data	3437
<i>Jianbin Qin, Wei Wang, Chuan Xiao, Ying Zhang</i>	
Building High Throughput Permissioned Blockchain Fabrics: Challenges and Opportunities	3441
<i>Suyash Gupta, Jelle Hellings, Sajjad Rahnama, Mohammad Sadoghi</i>	
Fairly Evaluating and Scoring Items in a Data Set.....	3445
<i>Abolfazl Asudeh, H. Jagadish</i>	

Award Talks, Keynotes, and Panel

Spitz: A Verifiable Database System	3449
<i>Meihui Zhang, Zhongle Xie, Cong Yue, Ziyue Zhong</i>	
Dremel: A Decade of Interactive SQL Analysis at Web Scale	3461
<i>Sergey Melnik, Andrey Gubarev, Jing Jing Long, Geoffrey Romer, Shiva Shivakumar, Matt Tolton, Theo Vassilakis, Hossein Ahmadi, Dan Delorey, Slava Min, Moshé Pasumansky, Jeff Shute</i>	
JIT works: decide when all data is known	3473
<i>Anastasia Ailamaki</i>	
Responsible Data Management	3474
<i>Julia Stoyanovich, Bill Howe, H.V. Jagadish</i>	
Out-of-order Execution of Database Queries	3489
<i>Kazuo Goda, Yuto Hayamizu, Hiroyuki Yamada, Masaru Kitsuregawa</i>	
The Relational Data Borg is Learning	3502
<i>Dan Olteanu</i>	
Winds from Seattle: Database Research Directions	3516
<i>Peter Bailis, Magda Balazinska, Xin Luna Dong, Juliana Freire, Raghu Ramakrishnan, Michael Stonebraker, Joseph M. Hellerstein</i>	