## The Third IEEE International Workshop on Benchmarking, Performance Tuning and Optimization for Big Data Applications (BPOD), Wednesday, December 11

Workshop Chairs: Zhiyuan Chen, Jianwu Wang, Feng Chen, Yiming Ying Workshop website: https://userpages.umbc.edu/~jianwu/BPOD

Session 1: Benchmark		worksnop website: <u>https://diserpages.umbc.ed</u>	w - Jianwa Di OD	
10:05-10:20	Time	Title	Presenter/Authors	
Multidimensional Preference Query Optimization on Infrastructure Monitoring Systems   Yinghua Qin and Gheorghi Guzun		Session 1: Benchmark		
10:40-11:00 Performance Benchmarking of Data Argumentation and Deep Learning for Tornado Prediction 11:00-11:20 Mix and Rank: A Framework for Benchmarking Recommender Systems 11:20-11:40 GeoYCSB: A Benchmark Framework for the Performance and Scalability Evaluation of NoSQL Databases for Geospatial Workloads 11:40-11:55 Towards a High-Level Description for Generating Stream Processing Benchmark Applications 11:55-12:10 Measuring, Quantifying, and Predicting the Cost-Accuracy Tradeoff 12:10-2:30 Lunch and Main Conference Keynote  Session 2: Analytics  2:30-2:50 An Empirical Study of Quad-Level Cell (QLC) NAND Flash SSDs for Big Data Applications 12:50-3:10 An Experimental Comparison of GPU Techniques for DBSCAN Clustering 3:10-3:30 A Gray-box Testing Method for Divide & Conquer in Image Processing 3:30-3:50 GHOSTZ PW/GF: distributed parallel homology search system for large-scale metagenomic analysis 3:50-4:10 Search for K: Assessing Five Topic Modeling Approaches to 120,000 Canadian Articles  Omega Target Ghosbal, Kesheng Wu, Eric Pouwaul, and Fric Pouwaul, a	10:05-10:20	An Empirical Study of Rabin Fingerprinting Parameters	Paul Lu, Owen Randall, and Emma McDonald	
11:00-11:20  Mix and Rank: A Framework for Benchmarking Recommender Systems  Bibek Paudel, Dragi Kocev, and Tome Eftimov  Systems  Bibek Paudel, Dragi Kocev, and Tome Eftimov  Scalability Evaluation of NoSQL Databases for Geospatial Workloads  Towards a High-Level Description for Generating Stream Processing Benchmark Applications  Towards a High-Level Description for Generating Stream Processing Benchmark Applications  Measuring, Quantifying, and Predicting the Cost-Accuracy Tradeoff  Lunch and Main Conference Keynote  Session 2: Analytics  An Empirical Study of Quad-Level Cell (QLC) NAND Flash SSDs for Big Data Applications  An Experimental Comparison of GPU Techniques for DBSCAN Clustering  An Experimental Comparison of GPU Techniques for DBSCAN Clustering  3:10-3:30  A Gray-box Testing Method for Divide & Conquer in Image Processing  GHOSTZ PW/GF: distributed parallel homology search system for large-scale metagenomic analysis  3:50-4:10  Search for K: Assessing Five Topic Modeling Approaches to 120,000  Canadian Articles  Coffee Break  Session 3: Tuning  Devarshi Ghosbal, Kesheng Wu, Eric Pouvaul and Frie	10:20-10:40		Yinghua Qin and Gheorghi Guzun	
11:20-11:40  GeoYCSB: A Benchmark Framework for the Performance and Scalability Evaluation of NoSQL Databases for Geospatial Workloads  11:40-11:55  Towards a High-Level Description for Generating Stream Processing Benchmark Applications  Measuring, Quantifying, and Predicting the Cost-Accuracy Tradeoff  Lunch and Main Conference Keynote  Session 2: Analytics  An Empirical Study of Quad-Level Cell (QLC) NAND Flash SSDs for Big Data Applications  An Experimental Comparison of GPU Techniques for DBSCAN Clustering  An Gray-box Testing Method for Divide & Conquer in Image Processing  3:30-3:30  GHOSTZ PW/GF: distributed parallel homology search system for large-scale metagenomic analysis  Search for K: Assessing Five Topic Modeling Approaches to 120,000  4:10-4:30  Coffee Break  Search Gray-box Fediction of Data Transfer Throughput for Data.  Devarshi Ghoshal Kesheng Wu, Eric Pouvoul and Eric	10:40-11:00		Carlos Barajas, Matthias Gobbert, and Jianwu Wang	
11:40-11:55 Towards a High-Level Description for Generating Stream Processing Benchmark Applications  11:55-12:10 Measuring, Quantifying, and Predicting the Cost-Accuracy Tradeoff Linh Truong, and Kyle Chard  12:10-2:30 Lunch and Main Conference Keynote  Session 2: Analytics  An Empirical Study of Quad-Level Cell (QLC) NAND Flash SSDs for Big Data Applications  An Experimental Comparison of GPU Techniques for DBSCAN Clustering  A Gray-box Testing Method for Divide & Conquer in Image Processing  A GHOSTZ PW/GF: distributed parallel homology search system for large-scale metagenomic analysis  GHOSTZ PW/GF: distributed parallel homology search system for Canadian Articles  Coffee Break  Session 3: Tuning  Analysis and Prediction of Data Transfer Throughput for Data.  Devarshi Ghoshal Kesheng Wu, Eric Pouvoul and Eric	11:00-11:20		Bibek Paudel, Dragi Kocev, and Tome Eftimov	
Benchmark Applications  Keller  Matt Baughman, Nifesh Chakubaji, Krists Kreics, Hong Linh Truong, and Kyle Chard  Lunch and Main Conference Keynote  Session 2: Analytics  2:30-2:50  An Empirical Study of Quad-Level Cell (QLC) NAND Flash SSDs for Big Data Applications  An Experimental Comparison of GPU Techniques for DBSCAN Clustering  An Experimental Comparison of GPU Techniques for DBSCAN Clustering  A Gray-box Testing Method for Divide & Conquer in Image Processing  GHOSTZ PW/GF: distributed parallel homology search system for large-scale metagenomic analysis  Search for K: Assessing Five Topic Modeling Approaches to 120,000 Qiang Fu, Yufan Zhuang, Jiaxin Gu, Yushu Zhu, Huihn Canadian Articles  Session 3: Tuning  Analysis and Prediction of Data Transfer Throughput for Data.  Devershi Ghosbal Kesheng Wu, Frie Pouvoul and Frie	11:20-11:40		Suneuy Kim and Yuvraj Kanwar	
12:10-2:30  Lunch and Main Conference Keynote  Session 2: Analytics  2:30-2:50  An Empirical Study of Quad-Level Cell (QLC) NAND Flash SSDs for Big Data Applications  An Experimental Comparison of GPU Techniques for DBSCAN Clustering  An Gray-box Testing Method for Divide & Conquer in Image Processing  3:30-3:50  GHOSTZ PW/GF: distributed parallel homology search system for large-scale metagenomic analysis  3:50-4:10  Coffee Break  Session 3: Tuning  Linh Truong, and Kyle Chard  Shuwen Liang, Zhi Qiao, Sihai Tang, Song Fu, and Weisong Shi  Hamza Mustafa, Eleazar Leal, and Le Gruenwald  Marco Strutz, Hermann Hessling, and Achim Streit  Kenta Machida and Osamu Tatebe  3:50-4:10  Coffee Break  Session 3: Tuning	11:40-11:55		Alessio Pagliari, Fabrice Huet, and Guillaume Urvoy- Keller	
2:30-2:50 An Empirical Study of Quad-Level Cell (QLC) NAND Flash SSDs for Big Data Applications  2:50-3:10 An Experimental Comparison of GPU Techniques for DBSCAN Clustering  An Gray-box Testing Method for Divide & Conquer in Image Processing  GHOSTZ PW/GF: distributed parallel homology search system for large-scale metagenomic analysis  3:50-4:10 Search for K: Assessing Five Topic Modeling Approaches to 120,000 Qiang Fu, Yufan Zhuang, Jiaxin Gu, Yushu Zhu, Huihn Qin, and Xin Guo  Coffee Break  Session 3: Tuning  Analysis and Prediction of Data Transfer Throughput for Data.  Devarshi Ghoshal Kesheng Wu, Eric Pouvoul and Eric	11:55-12:10	Measuring, Quantifying, and Predicting the Cost-Accuracy Tradeoff	Matt Baughman, Nifesh Chakubaji, Krists Kreics, Hong- Linh Truong, and Kyle Chard	
2:30-2:50  An Empirical Study of Quad-Level Cell (QLC) NAND Flash SSDs for Big Data Applications  An Experimental Comparison of GPU Techniques for DBSCAN Clustering  A Gray-box Testing Method for Divide & Conquer in Image Processing  GHOSTZ PW/GF: distributed parallel homology search system for large-scale metagenomic analysis  Search for K: Assessing Five Topic Modeling Approaches to 120,000  General Search Sea	12:10-2:30			
Big Data Applications  Weisong Shi  2:50-3:10  An Experimental Comparison of GPU Techniques for DBSCAN Clustering  A Gray-box Testing Method for Divide & Conquer in Image Processing  GHOSTZ PW/GF: distributed parallel homology search system for large-scale metagenomic analysis  Search for K: Assessing Five Topic Modeling Approaches to 120,000 Canadian Articles  Coffee Break  Analysis and Prediction of Data Transfer Throughput for Data.  Devarshi Ghoshal Kesheng Wu, Eric Pouvoul and Eric				
Clustering  A Gray-box Testing Method for Divide & Conquer in Image Processing  GHOSTZ PW/GF: distributed parallel homology search system for large-scale metagenomic analysis  Search for K: Assessing Five Topic Modeling Approaches to 120,000 Qiang Fu, Yufan Zhuang, Jiaxin Gu, Yushu Zhu, Huihi Qin, and Xin Guo  Coffee Break  Session 3: Tuning  Analysis and Prediction of Data Transfer Throughput for Data.  Devarshi Ghoshal Kesheng Wu, Eric Pouvoul, and Eric	2:30-2:50			
3:30-3:50  GHOSTZ PW/GF: distributed parallel homology search system for large-scale metagenomic analysis  Search for K: Assessing Five Topic Modeling Approaches to 120,000 Canadian Articles  Coffee Break  Session 3: Tuning  Analysis and Prediction of Data Transfer Throughput for Data.  Devarshi Ghoshal Kesheng Wu, Eric Pouvoul, and Eric	2:50-3:10		Hamza Mustafa, Eleazar Leal, and Le Gruenwald	
3:50-4:10  Search for K: Assessing Five Topic Modeling Approaches to 120,000 Canadian Articles  Coffee Break  Session 3: Tuning  Analysis and Prediction of Data Transfer Throughput for Data.  Devarshi Ghoshal Kesheng Wu, Eric Pouvoul, and Eric	3:10-3:30		Marco Strutz, Hermann Hessling, and Achim Streit	
Canadian Articles  Qin, and Xin Guo  4:10-4:30  Coffee Break  Session 3: Tuning  Analysis and Prediction of Data Transfer Throughput for Data.  Devarshi Ghoshal Kesheng Wu, Eric Pouvoul, and Eric	3:30-3:50		Kenta Machida and Osamu Tatebe	
Session 3: Tuning  Analysis and Prediction of Data Transfer Throughput for Data.  Devarchi Ghoshal Kesheng Wu, Eric Pouvoul, and Eric	3:50-4:10		Qiang Fu, Yufan Zhuang, Jiaxin Gu, Yushu Zhu, Huihui Qin, and Xin Guo	
Analysis and Prediction of Data Transfer Throughput for Data.  Devarchi Ghoshal Kesheng Wu, Eric Pouvoul, and Eric	4:10-4:30	Coffee Break		
Analysis and Prediction of Data Transfer Throughput for Data-		Session 3: Tuning		
4:30-4:50 Intensive Workloads Strohmaier	4:30-4:50	Analysis and Prediction of Data Transfer Throughput for Data- Intensive Workloads	Devarshi Ghoshal, Kesheng Wu, Eric Pouyoul, and Erich Strohmaier	
4:50-5:10 Fast Stochastic Block Partitioning using a Single Commodity Machine Md Abdul Motaleb Faysal and Shaikh Arifuzzaman	4:50-5:10	Fast Stochastic Block Partitioning using a Single Commodity Machine	Md Abdul Motaleb Faysal and Shaikh Arifuzzaman	
5:10-5:30 Cluster-size optimization within a cloud-based ETL framework for Big Data Eftim Zdravevski, Petre Lameski, Ace Dimitrievski, Marek Grzegorowski, and Cas Apanowicz	5:10-5:30			
5:30-5:50 GraphOpt: a Framework for Automatic Parameters Tuning of Graph Processing Frameworks Muaz Twaty and Amine Ghrab	5:30-5:50		Muaz Twaty and Amine Ghrab	
5:50-6:10 Benchmarking the Discretisation Level of Continuous Attributes: Wanghu Chen, Chao Wang, Jing Li, Bo Yang, Yang Li and Jianwu Wang	5:50-6:10		Wanghu Chen, Chao Wang, Jing Li, Bo Yang, Yang Liu, and Jianwu Wang	
6:10-6:25 Reactive Microservices in Commodity Resources Divya Goel and Amaresh Nayak	6:10-6:25	Reactive Microservices in Commodity Resources	Divya Goel and Amaresh Nayak	