



Contact:

Kim Daugherty
Pervasive Software
512.231.6507

kimberli.daugherty@pervasive.com

For Release 7:00 a.m. CDT

Pervasive Software to Present Groundbreaking High-Performance, Massively Parallel Data Mining Results at KDD-09

Pervasive DataRush Scales Hundreds of Times Faster than Traditional Methods

AUSTIN, Texas – May 28, 2009 – [Pervasive Software® Inc. \(NASDAQ: PVSW\)](#), a global value leader in [data management](#) software, announced today at KDD-09 it will demonstrate a data mining engine that, based on internal benchmarks, is dramatically faster and more scalable than current alternatives. Its innovative parallel data mining efforts will be spotlighted at the 15th Annual Conference on Knowledge Discovery and Data Mining (KDD-09) June 28-July 1 in Paris, France.

“One of the most interesting benchmarks today is the Netflix dataset, with more than 100 million movie reviews created by more than 480,000 users and covering a set of more than 17,000 movies,” said Pervasive Innovation Labs Director Nena Marín, Ph.D. “In the process of experimentation and research on parallel data mining, we discovered compelling results with Pervasive DataRush. We ran k -means clustering algorithm on the entire Netflix dataset with $k=30$ in just 17 seconds, and significantly, that was on a \$2,500 commodity 8-core server.”

“Thanks to Pervasive DataRush’s dataflow technology and inherent dynamic scalability – it runs faster as you add more cores,” said Mike Hoskins, Pervasive CTO. “We enjoy near-linear scaling for both dataset size and runtimes that others in the industry can’t touch. For example, we loaded the same Netflix data into a popular data mining tool and ran k -means on the data. Because of the other tool’s memory-bound legacy architecture, it could only load 0.1 percent of the Netflix data. Even on that small data subset, the runtime was 90 minutes – versus Pervasive DataRush-powered processing of 100 percent of the data in 17 seconds. I’m convinced we could tackle datasets hundreds of times the size of the Netflix data on the same commodity hardware server.”

In the Pervasive Innovation Labs, Pervasive DataRush was also used to tackle the telecoms dataset from the 2009 KDD Cup challenge. The highly dimensional, very large and complex 15,000-column dataset required substantial data preparation and conditioning prior to training, model-building and testing exercises. All the preprocessing steps were executed across dozens of gigabytes of original and staging datasets by Pervasive DataRush in minutes on a commodity multicore server. Meanwhile, the incumbent solution based on a leading tool reportedly ran for 12 hours to execute a subset of the same preprocessing stages on the telecoms data.

In a third recent case in collaboration with a leading academic institution, Pervasive deployed Pervasive DataRush to model customer behavior for a different, extended recommender system with content-based and collaborative filtering. “In comparison with a leading 32-bit tool, Pervasive DataRush trained the model in just three seconds versus seven minutes with the alternative,” Marín noted.

“Pervasive DataRush represents a genuine breakthrough in high-volume, large-scale, data-intensive analytics and processing,” said Hoskins. “Pervasive DataRush brings a degree of fine-grained massive parallelism not available elsewhere. It is designed to fully exploit the multicore revolution, and we are extremely pleased at the degree of off-the-chart performance and scalability that we can deliver on ubiquitous multicore commodity servers. I see this as a game-changer in data mining and analytics. I’m particularly pleased with its ‘green’ aspect: we enable breakthrough results on a single commodity server as an alternative to power-hungry cluster technology.”

Marín will present at KDD a paper titled “Pervasive Parallelism in Data Mining: Dataflow Solution to Co-clustering Large and Sparse Netflix Data,” co-authored with The University of Texas’ Professor Joydeep Ghosh, Schlumberger Centennial Chair Professor of Electrical and Computer Engineering at the University of Texas, Austin. The paper, selected from among 686 total submissions for presentation at the conference, details work to deliver performance improvements in the Netflix recommender system running a computationally intensive co-clustering algorithm.

Also at the conference, Hoskins will participate on a panel discussion titled “Open Standards and Cloud Computing,” moderated by Michael Zeller, CEO and co-founder of Zementis, Inc. The panel will take place at 2:00 p.m. on Tuesday, June 30.

About Pervasive Software

Pervasive Software (NASDAQ: PVSW) helps companies get the most out of their data investments through embeddable data management, agile data integration software and revolutionary next generation analytics. The embeddable Pervasive PSQL™ database engine allows organizations to successfully embrace new technologies while maintaining application compatibility and robust database reliability in a near-zero database administration environment. Pervasive's multi-purpose data integration platform accelerates the sharing of information between multiple data stores, applications, and hosted business systems and allows customers to re-use the same software for diverse integration scenarios. Pervasive DataRush™ is an embeddable high-performance software platform for data-intensive processing applications such as claims processing, risk analysis, fraud detection, data mining, predictive analytics, sales optimization and marketing analytics. For more than two decades, Pervasive products have delivered value to tens of thousands of customers in more than 150 countries with a compelling combination of performance, flexibility, reliability and low total cost of ownership. Through Pervasive Innovation Labs, the company also invests in exploring and creating cutting edge solutions for the toughest data analysis and data delivery challenges. Robin Bloor, founder of Bloor Research and partner at Hurwitz and Associates recently cited Pervasive as one of the [*10 IT Companies to Watch in 2009*](#). For additional information, go to www.pervasive.com.

Cautionary Statement

This release may contain forward-looking statements, which are made pursuant to the safe harbor provisions of the Private Securities Litigation Reform Act of 1995. All forward-looking statements included in this document are based upon information available to Pervasive as of the date hereof, and Pervasive assumes no obligation to update any such forward-looking statement.

###

All Pervasive brand and product names are trademarks or registered trademarks of Pervasive Software Inc. in the United States and other countries. All other marks are the property of their respective owners.