Computer Science 2.0: A New World of Data Management

Michael L. Brodie
Chief Scientist
Verizon Communications, USA
Michael.Brodie@verizon.com

ABSTRACT

Data management, one of the most successful software technologies, is the bedrock of almost all business, government, and scientific activities, worldwide. Data management continues to grow, more than doubling in data and transaction volumes every two years with a growth in deployments to fuel a $15 billion market. A continuous stream of innovations in capabilities, robustness, and features lead to new, previously infeasible data-intensive applications. Yet forty years of DBMS innovation are pushing DBMSs beyond the complexity barrier where one-size-fits-all DBMSs do not meet the requirements of emerging applications. While data management growth will continue based primarily on the relational data model and conventional DBMSs, a much larger and more challenging data management world is emerging. In the 1990's data under DBMS management reached 10% of the world's data. The six-fold growth of non-relational data in the period 2006–2010 will reduce that number to well below 5%.

We are entering the next generation of computing with a fundamentally different computing model and paradigm characterized technologically by multi-core architectures, virtualization, service-oriented computing, and the semantic web. Computer Science 2.0 will mark the end of the Computing Era with its focus on technology and the beginning of the Problem Solving Era with its focus on higher levels of abstraction and automation (i.e., intelligent) tools for real world (i.e., imprecise) domains in which approximate and ever-changing answers are the norm. This confluence of limitations of conventional DBMSs, the explosive growth of previously unimagined applications and data, and the genuine need for problem solving will result in a new world of data management.

The data management world should embrace these opportunities and provide leadership for data management in Computer Science 2.0. Two emerging areas that lack such guidance are service-oriented computing and the semantic web. While concepts and standards are evolving for data management in service-oriented architectures, data services or data virtualization has not been a focus of the DBMS research or products communities. Missing this opportunity will be worse than missing the Internet. The semantic web will become the means by which information is accessed and managed with modest projections of 40 billion pages with hundreds of triples per page - the largest distributed system in the world - the only one. Tim Berners-Lee and the World Wide Web Consortium view databases are being nodes that need to be turned inside out. What does the database community think? Semantic web services will constitute a programming model of Computer Science 2.0. How does data management fit into this semantically rich environment?

Computer Science 2.0 offers the data management community one of the biggest challenges in its forty-year history and opens up a new world of data management. Key to success in this new world will be collaboration with other disciplines whether at the technical level - partnering with the semantic technologies community to augment their reasoning capabilities with systems support - or at the problem solving level - partnering with real world domains as proposed by the new discipline of Web Science.

Biography of the Speaker

Dr. Michael L. Brodie is the Chief Scientist of Verizon Services Operations in Verizon Communications, one of the world's leading providers of communications services. Dr. Brodie works on large-scale strategic Information Technology (IT) opportunities and challenges for Verizon Communications senior executives. His primary interest is in delivering business value from advanced and emerging technologies and practices to enable business objectives while optimizing and transforming IT. In addition to Computer Science he addresses business and economic issues such as computing-communications-entertainment-business convergence and the economic impact of IT. His research focus is on advanced computational models and architectures and the large-scale information systems that they must support. He is concerned with the Big Picture, business and technical contexts, core technologies, and integration within a large scale, operational telecommunications environment.

Dr. Brodie holds a PhD in Databases from the University of Toronto.