NSF PI Meeting Application
"The Science of Cloud Computing"
March 17-18, 2011

Applicant:

Michael J. Carey, Computer Science Department, UC Irvine
(mjcarey@ics.uci.edu)

Two Primary Topics of Interest:

1. Cloud Architectures and Systems
3. Data Portability, Consistency, and Management

Cloud Computing Research Activities:

I am currently working on an NSF-funded project (with additional support from eBay, Facebook, and Yahoo!) entitled *ASTERIX: A Highly Scalable Parallel Platform for Semi-structured Data Management and Analysis*. The project's home base is UC Irvine, but it spans three UC campuses: UC Irvine, UC San Diego, and UC Riverside.

In the ASTERIX project, we are developing new technologies for ingesting, storing, managing, indexing, querying, analyzing, and subscribing to vast quantities of semi-structured information. The project is combining ideas from three distinct areas – semi-structured data, parallel databases, and data-intensive computing – to create a next-generation, open-source software platform that scales by running on large, shared-nothing commodity computing clusters. ASTERIX targets a wide range of semi-structured information, ranging from "data" use cases – where information is well-tagged and highly regular – to "content" use cases – where data is irregular and much of each datum is textual. ASTERIX is taking an open stance on data formats and addressing research issues including highly scalable data storage and indexing, semi-structured query processing on very large clusters, and merging parallel database techniques with today's data-intensive computing techniques to support performant yet declarative solutions to the analysis of semi-structured information.

As a foundation for the overall ASTERIX system, we are also developing Hyracks, a generalized alternative to infrastructures such as MapReduce (Hadoop) and Dryad for solving large-scale data-intensive computing problems. Hyracks balances a need for expressiveness beyond the MapReduce computational model with out-of-the-box support (unlike Dryad) for commonly occurring communication patterns and operators for data-oriented tasks. Hyracks is bringing efficient runtime techniques from parallel databases to bear on the problems in this space while aiming to deliver scalability and resiliency properties similar to those of Hadoop. Like ASTERIX, the Hyracks platform will be delivered as an open-source platform and made available for use by interested members of the data-intensive computing community.

Research Challenges

This is an exciting time to be working in data-intensive computing, as there is a wide range of research problems waiting to be tackled by our community. Problems of current and/or future interest to my UCI research agenda include:

1. *Scheduling and Fault-Tolerant Execution in Large Clusters.* Computing clusters found at large "cloud companies" typically involve many hundreds, thousands, or even tens of thousands of nodes interconnected in a (hierarchical) shared-nothing manner. Effective utilization of such large systems is currently an open problem. Hadoop takes an effective but expensive (pessimistic and simplistic) approach to the fault-tolerant execution of large jobs, and the scheduling control knobs and control algorithms for Hadoop clusters are manual and/or rudimentary in nature. Aims of my research group at UCI include the development of new, effective techniques for more selectively materializing intermediate results for use in fault-tolerance and for parallelizing and scheduling the operations involved in large jobs in ways that take their resource (I/O, CPU, memory, etc.) characteristics into account.

2. Data Management Architectures for the Cloud. Current cloud data management architectures have essentially grown up "by accident" outside the database systems community. For example, at Google, the layers start with an append-only file system (GFS) that provides an ordered byte stream abstraction for large data. On top of this layer sits a semi-structured table manager (BigTable), and then the software layers that provide declarative data storage and querying (e.g., Fusion Tables) sit above that layer. One can find similar layering in the open source community, with HDFS, key-value stores built on top of HDFS, and so on. In addition, current architectures such as key-value stores cut the end-to-end architecture at the data-selection level rather than the query level; this is known in the (old time, at least) database community to be a poor (very chatty!) cut point for data processing. A longer-term aim of my research group at UCI is to explore more appropriate layering of software abstractions and functionality, one that makes sense when one (re) considers cloud data management from first principles. It is our belief that building tomorrow's systems using components such as HDFS or Hadoop, or today's key-value stores, while expedient, is wrong-headed in the longer term, so we are "swimming against the current" and building cloud data management software from the ground ip.

In terms of research funding, my team is currently funded (from 2009-2012) by an NSF Large Grant from the Data-Intensive Computing cross-cutting program. Based on our progress in the first 1.3 years, we are optimistic about the research outcomes that our project should be able to yield. However, it is our desire to produce both Hyracks and ASTERIX as open-source code bases that can be utilized for actual applications in this space as well as being contributed to and further developed by others in our community. To that end, we are particularly interested in exploring possible funding sources and models that would enable us to staff the project in a manner that will enable us to provide sufficient robustness, documentation, and technical support to make that goal achievable.