Technical Perspective
Data Analysis at Astonishing Speed

By Michael J. Franklin

THE IMPORTANCE OF data analysis has never been clearer. Companies on the Web compete on their ability to intuit their users' wants, needs, and goals by analyzing their online behavior in ever-increasing volumes. Globe-spanning scientific collaborations are exploring data-intensive questions at a scale limited only by the power of the computing resources they can harness. The popular press regularly carries stories of analytics-driven applications as varied as timing the purchase of airline tickets, building a cost-efficient major league baseball team, and understanding social phenomena.

While data analytics and business intelligence have long been active areas of research and development, the past decade has seen a dramatic increase in the scale and complexity of the problems being addressed. Since its inception, Google has been aggressively pushing the envelope in terms of data management technology. Over the years they have published a series of highly influential papers describing key aspects of their internal systems.

A recent addition to this series is the following paper on the Dremel system. Dremel is an interactive data analysis tool capable of leveraging thousands of machines to process data at a scale that is simply jaw-dropping given the current state of the art in the field. The system enables Google engineers to analyze billion-record data sets interactively and trillion-record data sets records while they make a macchiatone at a nearby espresso machine.

Large Web properties track billions of events per day, representing tens of terabytes of data. Storing such data is expensive but doable; providing meaningful access to that data, however, is another story. The reality in most organizations today is analysis turnaround times measured in hours or even days. Such batch-oriented delays may have been tolerable for the more leisurely decision cycles of the past, but are severely limiting for today's fast pace of business and data acquisition. Furthermore, it is not simply a matter of latency. Moving from hours to seconds enables new modes of data exploration providing a qualitative change in the value that can be extracted from complex data sets.

Not surprisingly, Dremel's performance depends on massive parallelism. The authors point out that scanning 1TB in one second would require 10,000 commodity disk drives. But the main take-away from this paper is that simply throwing hardware at the problem is not sufficient. Rather, it is critical to deeply understand the structure of the data to be stored and how that data will be used.

Recent advances in database systems have exploited organization of data vertically along columns, rather than horizontally along rows, to reduce disk accesses through compression and selective retrieval. Dremel follows a similar philosophy but its task is complicated by the need to store irregular, nested data consisting of repeating fields, optional and variable fields, and structured sub-objects—all constructs that violate basic relational database principles.

Dremel describes an efficient, column-oriented structure for storing and sharding (that is, partitioning) such data, as well as a high-performance

> ## Dremel is an interactive data analysis tool capable of leveraging thousands of machines to process data at a scale that is simply jaw-dropping.

mechanism for reconstructing data records stored in such a fashion. A number of other important techniques are addressed, including an execution model based on serving trees, query-result caching, replication, and the willingness to return partial answers to avoid waiting for the "stragglers" (slow tasks) that inevitably arise in a massively parallel system.

The real eye-opener, however, is the experimental section. Prior to reading this paper, I would not have thought it possible to obtain the low latencies over massive data sets demonstrated here, with processing throughputs in the range of 100 billion records per second. The fact that Googlers have been using this technology already for five years was a revelation.

In this regard, the paper raises some serious questions for computing systems researchers. One could reasonably ask if such technology is relevant only to the very few organizations that currently have data at this huge scale. Similarly, some have raised the question of whether those outside of such organizations, without easy access to massive computing resources, can even attempt to participate in such research.

The answer to the first question is obvious. One only needs to look through past issues of this publication and others to see how quickly the "bleeding edge" becomes commonplace in our field. The data volumes described in this paper will clearly become relevant to more organizations over time. The second question is more difficult. Looking at the paper objectively, I see numerous opportunities for optimizations and improvements, such as stream-oriented processing and more aggressive sampling. Arguably, those without easy access to thousands of machines should be particularly motivated to explore such innovative techniques.

In summary, Dremel is a very impressive system that extends our idea of what is possible today and what will be required in the future. As such, this paper makes compelling reading for anyone interested in the future of data management. 🄲

Michael J. Franklin (franklin@cs.berkeley.edu) is a professor of computer science at the University of California, Berkeley and a founder and CTO of Truviso, Inc.