Timely Dataflow

In this book we will work through the motivation and technical details behind timely dataflow, which is both a system for implementing distributed streaming computation, and if you look at it right, a way to structure computation generally.

Timely dataflow arose from work at Microsoft Research, where a group of us worked on building scalable, distributed data processing platforms. Our experience was that other systems did not provide both *expressive computation* and *high performance*. Efficient systems would only let you write restricted programs, and expressive systems employed synchronous and otherwise inefficient execution.

Our goal was to provide a not-unpleasant experience where you could write sophisticated streaming computations (e.g. with iterative control flow), which nonetheless compile down to systems that execute with only a modicum of overhead and synchronization.