Sade: A Smooth Operator for Partitioning Data in the Background

ABSTRACT

Parallel databases must partition or *shard* data across multiple nodes, ideally in a dynamic fashion that adjusts to changes in data and workload distributions. This is typically done via a partitioning operator p mapping items to nodes. From calculus we have that p is a *smooth* function if it has continuous derivatives of all orders at all points $x \in \mathbb{R}$. Smooth functions are well suited to dynamic, fine-grained background tasks due to their infinite differentiability. We identify such a smooth operator, called **Sade**, for fine-grained dynamic data partitioning.

We present an extensive empirical evaluation of Sade on workloads that are geographically distributed across data centers in the cloud (Coast-to-Coast, L.A.-to-Chicago). We find that Sade is an excellent background task for a variety of applications including online shopping, sexting and remote dentistry. In elevator scheduling and telemarketing applications, however, we measure noticeably increased wait time perception, which we attribute to inappropriate use of Sade as a foreground process.

BODY

Infinitely differentiable (smooth) functions are well-suited to background tasks. Sade *is such a smooth operator for data partitioning.*

Volume X of Tiny Transactions on Computer Science