CS3550-ADMT25 Guests Lectures

Date: Mar 2, 2021

Title: Avoiding class warfare: Managing Continuous Queries with Differentiated Classes of Service

Abstract: Data stream management systems (DSMSs) offer the most effective solution for processing data streams by efficiently executing continuous queries (CQs) over the incoming data. CQs inherently have different levels of criticality and hence different levels of expected quality of service (QoS) and quality of data (QoD). Adhering to such expected QoS/QoD metrics is even more important in cases of multi-tenant data stream management services. In this work, we propose DILoS, a framework that, through priority-based scheduling and load shedding, supports differentiated QoS and QoD for multiple classes of CQs. Unlike existing works that consider scheduling and load shedding separately, DILoS is a novel unified framework that exploits the synergy between scheduling and load shedding. We also propose ALoMa, a general, adaptive load manager that DILoS is built upon. By its design, ALoMa performs better than the state-of-the-art alternatives in three dimensions: (1) it automatically tunes the headroom factor, (2) it honors the delay target, (3) it is applicable to complex query networks with shared operators. We implemented DILoS and ALoMa in our real DSMS prototype system (AQSIOS) and evaluate their performance for a variety of real and synthetic workloads. Our experimental evaluation of ALoMa verified its clear superiority over the state-of-the-art approaches. Our experimental evaluation of the DILoS framework showed that it (a) allows the scheduler and load shedder to consistently honor CQs' priorities, (b) significantly increases system capacity utilization by exploiting batch processing, and (c) enables operator sharing among query classes of different priorities while avoiding priority inversion, i.e., a lower-priority class never blocks a higher priority one.

Speaker: Dr. Thao N. Pham is an alumni of Pitt CS and the ADMT Lab. Her PhD thesis was on the topic of load shedding and scheduling in Data Stream Management Systems. Currently she is a senior systems/software engineer at Vertica, working on query optimization and distributed infrastructure.