Tuning, Optimizing, and Selecting the right way to access data

Manos Athanassoulis (Harvard U)

Host: Bernd Bickel

Business and science create an increasing amount of digital data every day; these ever growing data collections require analytical workflows that support fast ingestion, and fast read queries; in other words, data systems need to support increasingly heterogeneous workloads, that weave read-mostly analytical operations and update-heavy transactional operations, into hybrid transactional/analytical processing workloads. In this talk I will present a path towards building access methods for this new set of workloads. As the design goals typically entail to reduce read latency, update latency, and space utilization, I attack the problem by studying the trade-offs between these design goals, and I propose solutions that balance the trade-offs through modeling and optimization. I will use this methodology to build access methods for modern key-value stores, basic column storage, and to present how to perform access path selection in light of the changes in workloads and hardware. I will conclude with abstracting the common concepts of access method designs, and propose a path towards data access methods for the new era.