Chapter 9
Adaptive Query Processing in Distributed Settings

Anastasios Gounaris, Efthymia Tsamoura, and Yannis Manolopoulos

Abstract. In this survey chapter, we discuss adaptive query processing (AdQP) techniques for distributed environments. We also investigate the issues involved in extending AdQP techniques originally proposed for single-node processing so that they become applicable to multi-node environments as well. In order to make it easier for the reader to understand the similarities among the various proposals, we adopt a common framework, which decomposes the adaptivity loop into the monitoring, analysis, planning and actuation (or execution) phase. The main distributed AdQP techniques developed so far tend to differ significantly from their centralized counterparts, both in their objectives and in their focus. The objectives in distributed AdQP are more tailored to distributed settings, whereas more attention is paid to issues relating to the adaptivity cost, which is significant, especially when operators and data are moved over the network.

9.1 Introduction

The capability of database management systems to efficiently process queries, which are expressed as declarative statements, has played a major role in their success over the last decades. Efficiency is guaranteed due to several sophisticated optimization techniques, which heavily rely on the existence of metadata information about the data to be processed, such as the distribution of values and the selectivity of the relational operators. Since the late 1970s and the introduction of System R [58], static optimization of query plans and subsequent execution has been the main choice for database system developers. However, when the metadata required are not available or accurate at compile time, or when they change during execution,
the query processor needs to revise the current execution plan on the fly. In this case, query processing is called adaptive.

In adaptive query processing (AdQP), there is a feedback loop, similar to the one appearing in autonomic systems, according to which the query processor monitors its execution properties and its execution environment, analyzes this feedback, and possibly reacts to any changes identified with a view to ensuring that either the current execution plan is the most beneficial or a modification of the current plan can be found that is expected to result in better performance.

Although AdQP is particularly relevant to wide area settings, in which query statistics are more likely to be limited or potentially inaccurate, and the computational properties, such as the processing capacity of hosting machines, are volatile, most AdQP proposals have focused either on completely centralized query processing or on centralized processing of data retrieved or stemming from remote sources and data streams, respectively. In such settings, there is typically a single physical machine used for query execution, which is predefined, and thus the focus is mostly on adapting to changing properties of the data processed, e.g., cardinalities of intermediate results and operator selectivities. This is, of course, of high importance for distributed query processing (DQP), as crucial information about the data may be missing at compile time. However, of equal significance are adaptations to changing properties of a potentially arbitrary set of resources that DQP may employ and of their communication links. Currently, AdQP with respect to changing resources is not addressed as satisfactorily as with respect to changing data properties.

In this survey chapter, we systematically discuss AdQP techniques that are tailored to distributed settings both with respect to the data sources and the processing nodes. We also investigate the issues involved in extending AdQP techniques originally proposed for single-node processing so that they become applicable to multi-node environments as well. In order to make it easier for the reader to understand the similarities among the various proposals, we adopt a common framework, which decomposes the adaptivity loop into its constituent phases mentioned above, i.e., monitoring, analysis, planning and actuation phase. The later corresponds to the phase, in which the adaptivity decisions are executed by the system.

**Structure.** The structure of this chapter is as follows. In the remainder of this section we briefly discuss preliminary concepts of distributed query processing and optimization (Section 9.1.1), and related work (Section 9.1.2). In Section 9.2 we present the framework that forms the basis of our analysis. The next section contains a short review of traditional AdQP for centralized settings and explains the reasons why such techniques cannot be applied to wide-area environments in a straightforward manner. The discussion of the AdQP techniques for distributed settings, which is the core part of this chapter, is in Sections 9.4-9.6. Existing work in distributed AdQP techniques can be classified in three broad categories. Techniques that do not rely on the existence of traditional query plans fall into the first category, which is examined in Section 9.4. The second category comprises approaches that perform load management at the operator level (Section 9.5), whereas, in Section 9.6 we discuss distributed AdQP techniques where the adaptivity occurs at a higher level.