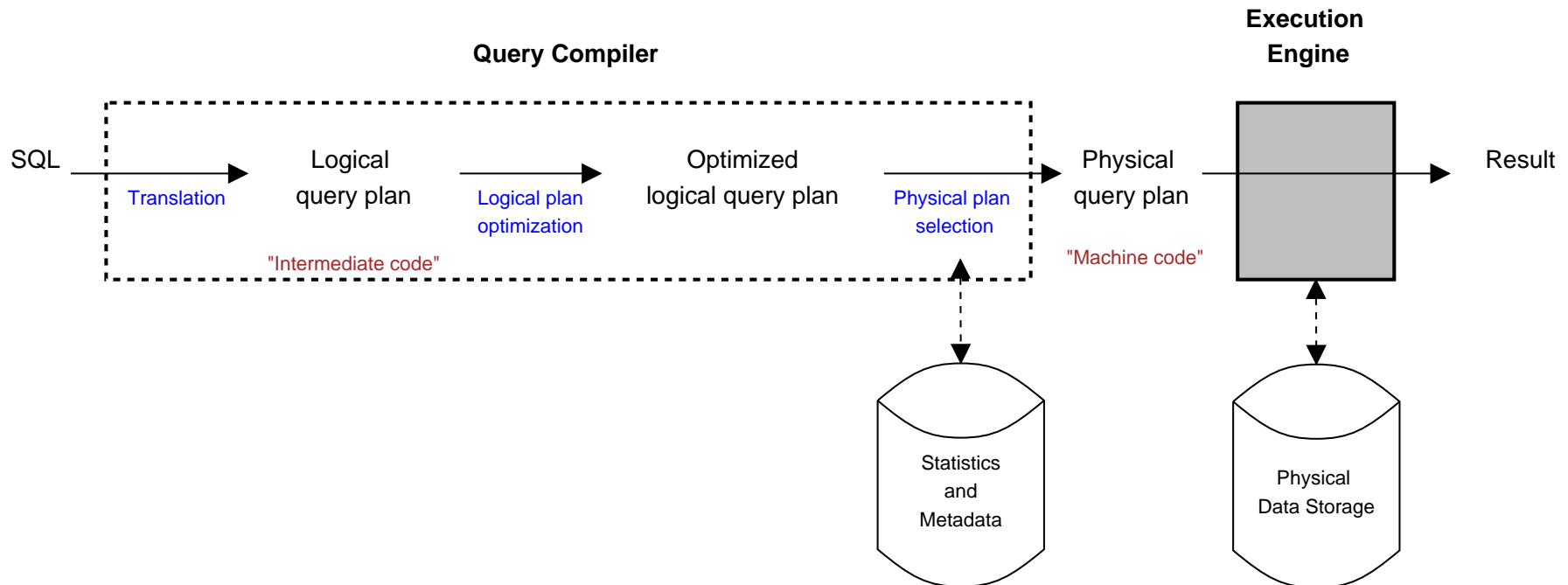


Physical data organization

Disks, blocks, tuples, schemas

Physical data organization



In order to select a physical plan we need to know:

- The physical algorithms available to implement the relational algebra operators
e.g., scan a relation to implement a selection
- The situations in which each algorithm is best applied (situation x calls for algorithm A , situation y calls for algorithm B , ...).

Physical data organization

Physical algorithms depend on

- The representation of data on disk
- The data structures used

We hence need to know how data is physically organized before studying algorithms

This is the subject of chapters 13 and 14 in the book

Physical data organization

Database Management System

- Are responsible for **enormous quantities** of data (current scale: exabytes = 1 million gigabytes)
- Must query this data as efficiently as possible
- Must store data as reliably as possible

Hence we should wonder:

- What are the available storage media?
- How much “time” does it take to read from/write to these media?
- How can we minimize this costs?
- How can we prevent data loss due to disk crashes?

The answers to these questions may be found in chapter 13

Physical data organization

The types of data that we will need to store are:

- Schemas
- Records
- Relations

How can we represent them efficiently “on disk”?

The answer may be found in chapter 13