

# Basic Approach



## Question

Is the observed performance acceptable?

## Basic Approach

Start with a simple model

- 1 Measure time for the execution of your workload
- 2 Quantify the workload with some metrics
  - ▶ E.g., amount of tuples or data processed, computational operations needed
  - ▶ E.g., you may use the statistics output for each Hadoop job
- 3 Compute  $w_t$ , the workload you process per time
- 4 Compare  $w_t$  with your expectations of the system

Refine the model as needed, e.g., include details about intermediate steps