Tutorial 7 Solutions

Solutions

1. a) The saving would amount to

$$0.25 * (0.13 + 0.18) = 0.0775,$$

so about 8% of the total costs.

b) The saving would amount to

$$0.25 * 0.57 = 0.1425,$$

so about 14% of the total costs.

c) The saving would amount to

$$0.0775 + 0.1425 = 0.22,$$

which is 22% of the total costs.

2. Figure 1 shows the setup of the idealized web search application whose index is distributed over N servers. Here we assume that all queries are answered independently and that servers are always idle when queries arrive. Hence, we can consider only a single query and determine the expected latency of its response.

Let R_i be the response time of the i-th server. The problem states that R_i is chosen via a Bernoulli trial as follows:

$$\begin{aligned} \text{Prob}(R_i = 2 \text{ ms}) &= p := 0.99 \\ \text{Prob}(R_i = 100 \text{ ms}) &= 1 - p = 0.01. \end{aligned}$$

Note that p is the "success probability", which is the probability of having a faster response time.

Let R be the response time of the complete web search machine. Note that a query can only be answered when the last server has sent its response, which means that

$$R = \max_i R_i, \quad 1 \le i \le N.$$

From this follows directly that

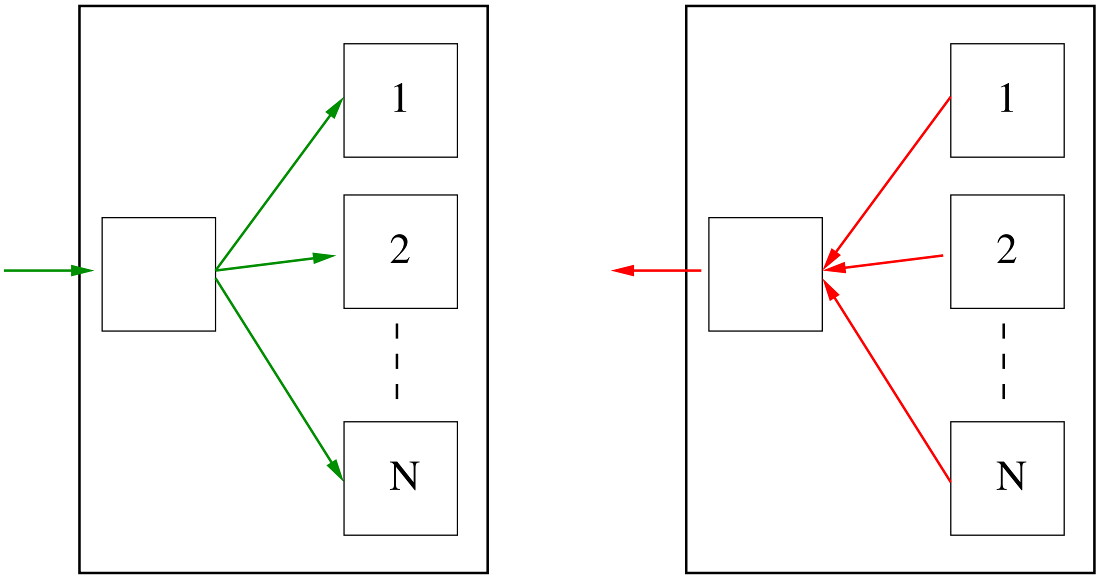
$$\text{Prob}(R = 2 \text{ ms}) = \prod_i \text{Prob}(R_i = 2 \text{ ms}) = p^N$$

and thus

$$\text{Prob}(R = 100 \text{ ms}) = 1 - \text{Prob}(R = 2 \text{ ms}) = 1 - p^N.$$

Therefore, we obtain for the expected value $E[R]$

$$E[R] = 2\,ms * \text{Prob}(R = 2 \text{ ms}) + 100\,ms * \text{Prob}(R = 100 \text{ ms}) = 2\,ms * p^N + 100\,ms * (1 - p^N).$$



(a) Arrival of search query (b) Return of search results

Figure 1: Visualization of a search machine whose index is partitioned into N parts, each served by a separate server.

a) For $N = 10$: $E[R] \approx 11.37\ ms$

b) For $N = 100$: $E[R] \approx 64.13\ ms$

c) For $N = 1000$: $E[R] \approx 100\ ms$