

Exam 2018 (January)

June 19, 2023

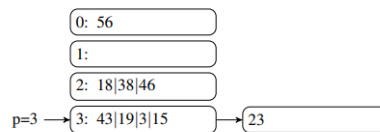
Exercise 0.1. Which two kinds of schema information contains the Global Conceptual Schema that does not contain the Local Conceptual Schema in the Extended ANSI-SPARC Architecture for DDBMS?

1. Fragmentation
2. Allocation

Exercise 0.2. If you have a stream of messages that arrive at a rate of 10.000 per second accounting for 512MB per hour, how many RDD will process your Spark Streaming process in one hour? Explicit all the assumptions you make.

It depends on our configuration. We can configure Spark to trigger every t seconds, so it will process $\frac{3600}{t}$ RDDs.

Exercise 0.3. Briefly explain what is wrong in this linear hash structure, or if you think it is right, explicitly say so:



In this case, $n = 2$, and it is not possible to have the pointer pointing to a bucket between 2^n and 2^{n+1} , because the design of the algorithm makes this impossible. When 2^n is reached, the pointer goes back to the first bucket and the process starts all over.