For the upcoming lesson on November 14, 2024, prepare the self-study material in advance and answer the questions under Q&A. Bring your answers to the next lesson.
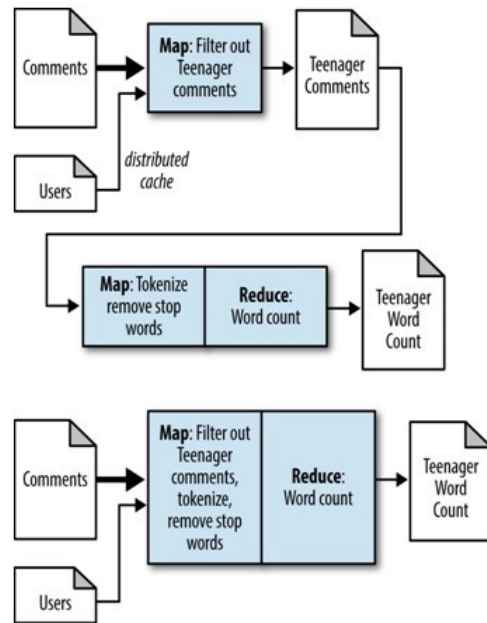
Self Study Material

This unit is based on the book chapter "Chain Folding" from the O'Reilly book "MapReduce Design Patterns" by Donald Miner and Adam Shook.

Q&A

1. What is the benefit of merging a two-job-chain of a Map-only job and a traditional MapReduce job?

What is a downside?

2. Explain this example of chain folding:

3. Why should you avoid to merge jobs that use a lot of main memory?