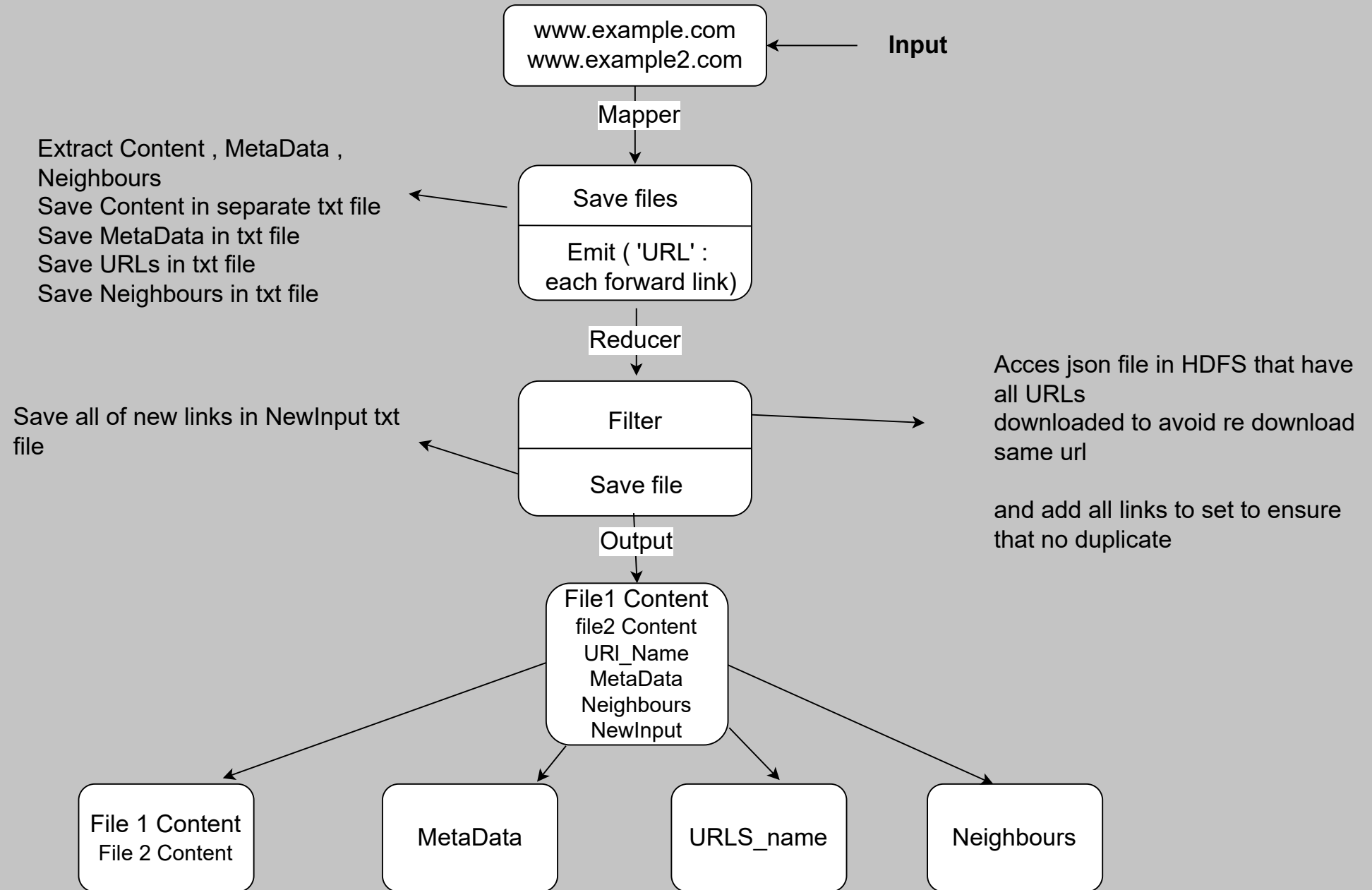


1-

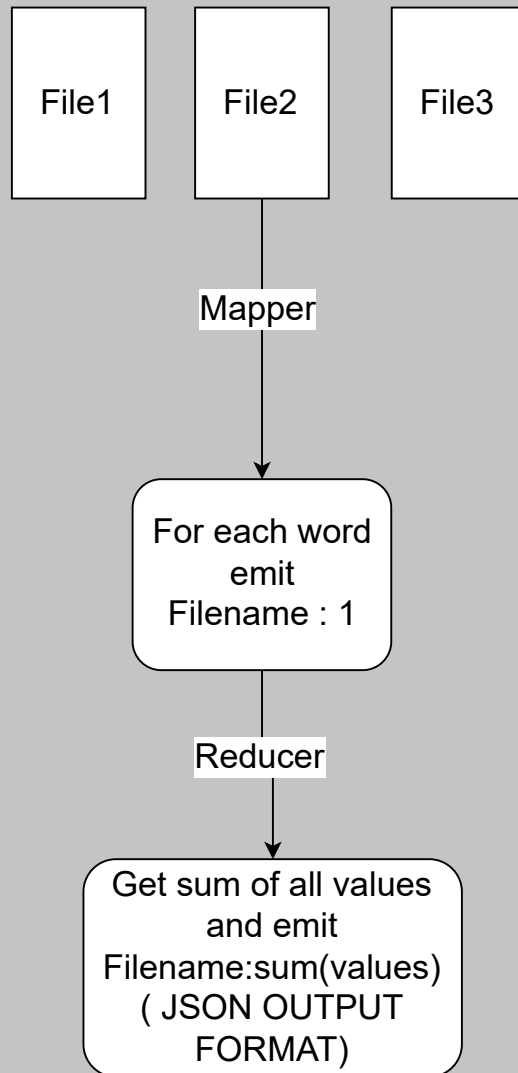
crawling with map reduce



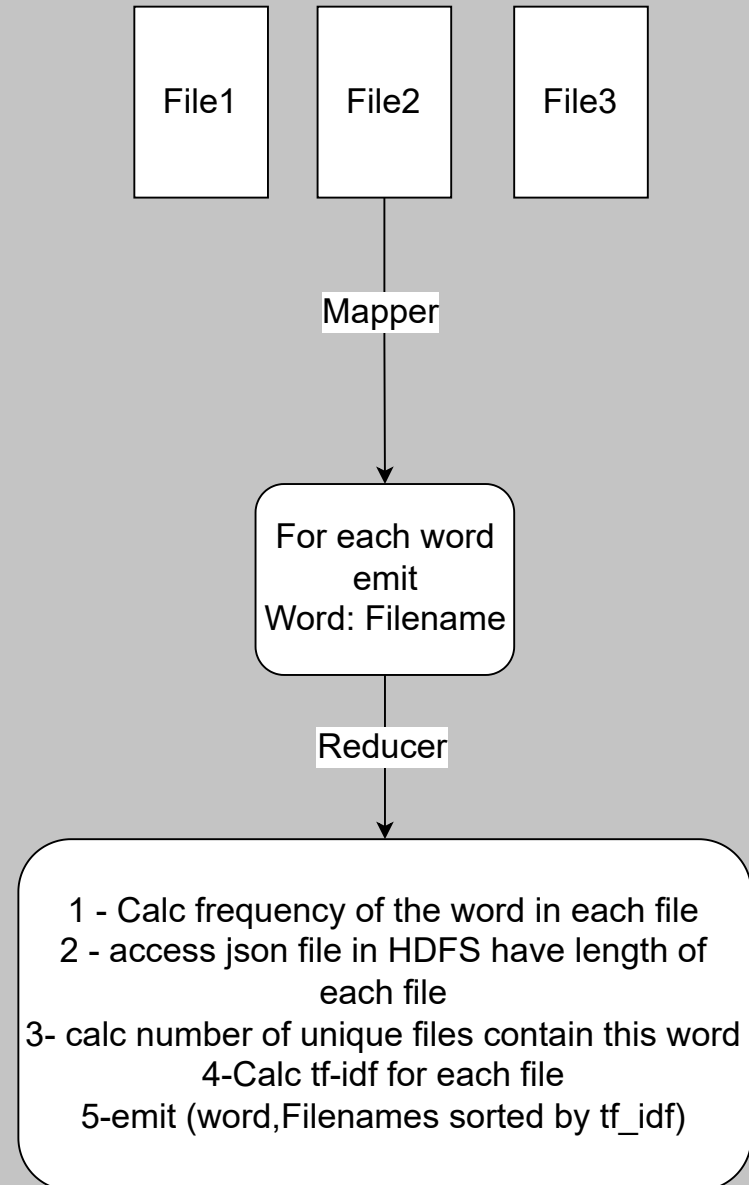
2-

- 3- replace seed points with NewInput file
- 4- Apply Map reduce (Convert txt to json) on URLS.txt
- 5 - goto step 1

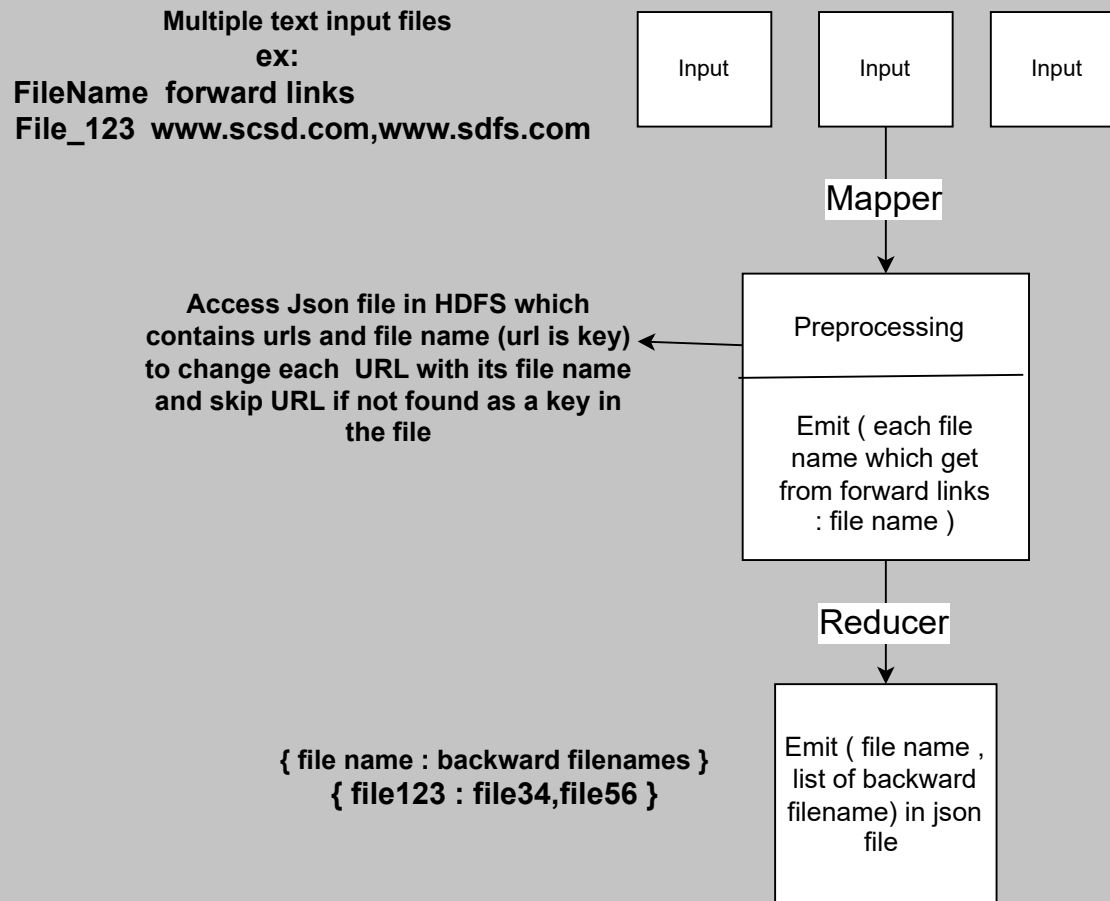
Calc file length with map reduce



Calc Inverted index and tf_idf with map reduce



Backward link generate Map reduce



Map reduce to convert txt files to Json

