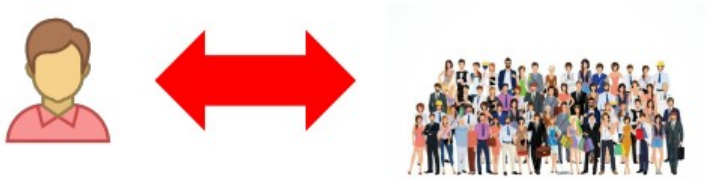
CORK INSTITUTE OF TECHNOLOGY
INSTITIÚID TEICNEOLAÍOCHTA CHORCAÍ

BIG DATA PROCESSING.

Lab02: Benchmark Generator & Solver (Lab Demonstration).

BACKGROUND.

The code examples associated to ***Lecture 03-05. Distributed Programming*** provide us with benchmark generator & solver app for running the problem of computing list inversions of person P vs an entire population.



This lab demonstration explains the benchmark generator & solver app in detail.

The code examples contain 3 folders:

- code:
  - Contains the Python code of the application.
- input_files:
  - Contains the benchmark files generated.
- results:
  - Contains a report with the results from running the benchmark.

PYTHON APPLICATION.

The folder code contains 4 files:

1. my_main.py
   - Main entry point to the application.
   - Function my_main: Triggering a benchmark generation and its solving.
   - Function get_report_name: Generating the name of the report.
   - Function write_report: Filling the report file with the benchmark running results.

2. create_benchmark.py
   - Function generate_file: Creating a single file of the benchmark.
   - Function generate_benchmark: Generating the entire benchmark.

3. solve_benchmark.py
   - Function run_file: Solving one single file from the benchmark.
   - Function my_divide_stage: Distributed approach – Divide Stage.
   - Function core_workload: Solving the files associated to 1 core by the Map Stage.
   - Function my_map_stage: Triggering the work of all cores in parallel.
   - Function my_reduce_stage: Merging the results of all cores.
   - Function run_benchmark: Solving all the benchmark.

4. problem_algorithms.py
   - Function count_inversions_n2: Computing list inversions via the $O(n^2)$ algorithm.
   - Function merge_inversions: Reduce stage of the O(n logn) algorithm.
   - Function mergesort_and_count_inversions: Divide-Map-Reduce stages of the O(n logn) algorithm.
   - Function count_inversions_n_logn: Computing list inversions via the O(n logn) algorithm and discarding the sorted array A' as it is irrelevant to us.

LAB DEMONSTRATION VIDEO.

Topic: 02. Lab02_22_09_2019
Date: Sep 22, 2019 06:06 PM Dublin

Meeting Recording:
https://telcit.zoom.us/recording/share/SjU9X7Otx1QdamXTBOol9cXLL4VW367GpnAJbVA2LzawIumekTziMw