

Experimental Setup

- We have 2 candidate algorithms. The exhaustive search and the A*-like algorithm.
- In all experiments in the vertical axis we measure the execution time (in milliseconds) of the candidate algorithms (1 plot).
- In all experiments in the vertical axis we count the number of plans that reach the OR operator for each of the candidate algorithms (a separate plot).

Experiments with Real Workflows (images of workflows per use case follow at the end of the document)

Network topology: any computer cluster/Big Data platform can communicate with others directly in 1 hop.

Experiment 1: Fixed 1 cluster with 3 Big Data platforms for all use cases. Assume 3 implementations exist for all operators of all workflows in all 3 Big Data platforms.

Result:

1 plot with execution time in vertical axis, 1 plot with number of plans reaching the OR operator in vertical axis

Horizontal axis → use case (Life Science, Financial, Maritime).

For each use case, each plot will have two bars (1 for exhaustive, 1 for A*)

Experiment 2: 1 cluster with 3 Big Data platforms for the Life Sciences use case, 9 computer clusters with 1 Big Data platform each for the Financial use case, 100 clusters for the Maritime use case with 1 Big Data platform each. Horizontal axis → use case (Life Science, Financial, Maritime).

Result:

1 plot with execution time in vertical axis, 1 plot with number of plans reaching the OR operator in vertical axis

Horizontal axis → use case (Life Science, Financial, Maritime).

For each use case, each plot will have two bars (1 for exhaustive, 1 for A*)

Experiment 3: Using the workflow of each use case. Run both algorithms varying the number of clusters between 2 and 10 (step 2). Each cluster hosts 2 Big Data platforms with implementations for all operators.

Result for each use case:

1 plot with execution time in vertical axis, 1 plot with number of plans reaching the OR operator in vertical axis

Horizontal axis → Number of clusters (2,4,6,8,10)

For each use case, each plot will have two lines (1 for exhaustive, 1 for A*)

6 plots in total for all use cases

Experiment 4: Using the workflow of each use case. Run both algorithms for fixed number of clusters=5. Each cluster hosts 1 to 5 (step 1) Big Data platforms with implementations for all operators.

Result for each use case:

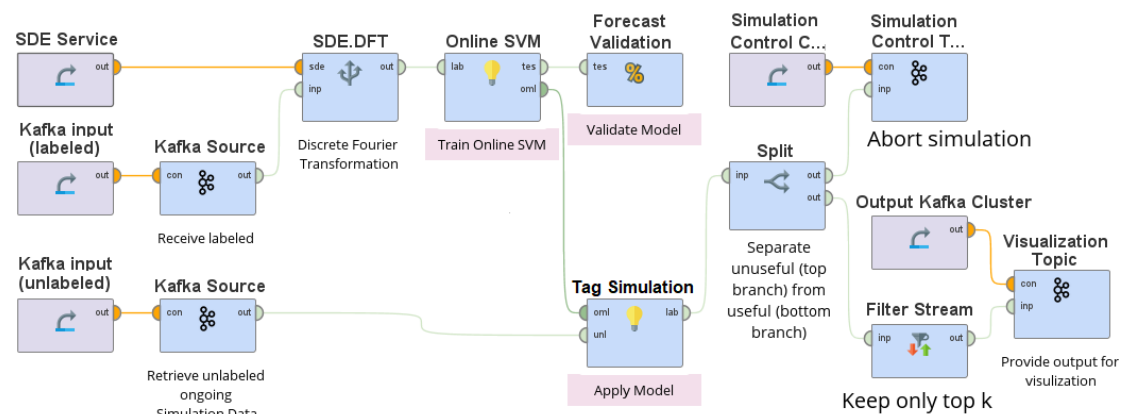
1 plot with execution time in vertical axis, 1 plot with number of plans reaching the OR operator in vertical axis

Horizontal axis → number of Big Data platforms (1,2,3,4,5)

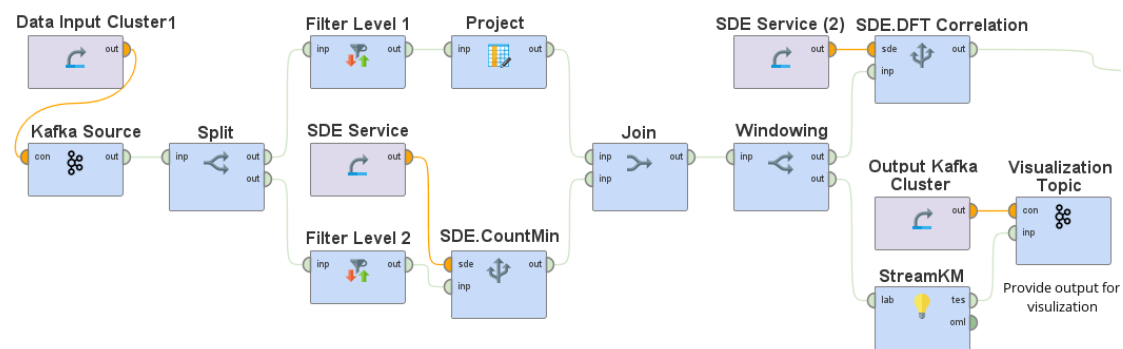
For each use case, plot will have two lines (1 for exhaustive, 1 for A*)

6 plots in total for all use cases

Workflow for life science use case:



Workflow for Financial use case:



Workflow for Maritime use case:

reaming Optimization

