

Installation Instructions for Parallel Trajectory Management in SECONDO

Fabio Valdés, Thomas Behr, and Ralf Hartmut Güting

October 31, 2019

1 Install Distributed Secondo

The SECONDO website is located at `dna.fernuni-hagen.de/secondo`. There are different options for using SECONDO, depending on the operating system and application, that are described in the following.

1.1 Virtual Machine

For basic applications with small datasets or for trying SECONDO without installation effort, it is preferable to apply the provided virtual machine appliance. On a Microsoft Windows machine, this is probably the only option for running SECONDO. The provided virtual machine file can be executed with VMware Workstation Player (<https://www.vmware.com/de/products/workstation-player.html>). Distributed features are already enabled in the SECONDO version available in the appliance.

1.2 Full Installation

1.2.1 Ubuntu, openSUSE

Initially, follow the provided installation instructions. In order to use SECONDO's distributed features, it is required to activate certain algebra modules.

As a precondition for the Distributed2Algebra to be executed, the Boost library has to be installed, either with the graphical package installation manager or in a terminal using one of the following commands:

- openSUSE: `sudo zypper in libboost*devel`
- Ubuntu: `sudo apt-get install libboost-all-dev`

After that, open the `makefile.algebras` file in SECONDO's main directory with a text editor. Remove the hashtag symbols in front of the following lines:

```
ALGEBRA_DIRS += Distributed2
ALGEBRAS += Distributed2Algebra
DEFAULTCCFLAGS += -pthread -DTHREAD_SAFE
CCFLAGS += -pthread -DTHREAD_SAFE
COMMON_LD_FLAGS += -lboost_thread -lboost_system

ALGEBRA_DIRS += DFS
ALGEBRAS += DFSAlgebra
COMMON_LD_FLAGS += -L$(SECONDO_BUILD_DIR)/Tools/DFS/dfs -ldfs
COMMON_LD_FLAGS += -L$(SECONDO_BUILD_DIR)/Tools/DFS/shared -lshared
COMMON_LD_FLAGS += -L$(SECONDO_BUILD_DIR)/Tools/DFS/commlayer -lcommlayer
```

After saving and exiting, the SECONDO system has to be built from scratch. In the main directory, execute `make clean; make`.

1.2.2 Mac OS-X

Initially, follow the provided installation instructions. For activating the Distributed2Algebra and its dependencies, first open the `makefile.algebras` file in SECONDO's main directory with a text editor. Remove the hashtag symbols in front of the following lines:

```
ALGEBRA_LINK_FLAGS += -framework CoreServices

ALGEBRA_DIRS += Distributed2
ALGEBRAS += Distributed2Algebra
DEFAULTCCFLAGS += -pthread -DTHREAD_SAFE
CCFLAGS += -pthread -DTHREAD_SAFE
COMMON_LD_FLAGS += -lboost_thread-mt -lboost_system-mt

ALGEBRA_DIRS += DFS
ALGEBRAS += DFSAlgebra
COMMON_LD_FLAGS += -L$(SECONDO_BUILD_DIR)/Tools/DFS/dfs -ldfs
COMMON_LD_FLAGS += -L$(SECONDO_BUILD_DIR)/Tools/DFS/shared -lshared
COMMON_LD_FLAGS += -L$(SECONDO_BUILD_DIR)/Tools/DFS/commlayer -lcommlayer
```

After saving and exiting, the SECONDO system has to be built from scratch. In the main directory, execute `make clean; make`.

1.2.3 Other Systems

For installing SECONDO on another system, please refer to the instructions for manual installation of all required tools and setting the environment variables. As for the mentioned systems, the Boost library has to be installed and the Distributed2Algebra has to be activated.

2 Start Local Workers

For easily starting several workers on a single machine (e.g., for testing or presentation purposes), download the bash script `dna.fernuni-hagen.de/secondo/files/startLocalWorkers` to the `secondo/bin` directory. The following command starts four local workers with ports 1300 to 1303 and creates a suitable worker relation in the file `workers.local`:

```
startLocalWorkers -c SecondoConfig.ini -f Workers.local 1300 4
```

3 Download and Prepare the Dataset

The dataset applied in the tutorial can be downloaded from <https://crawdad.org/roma/taxi/20140717/> (note that a free registration is required for accessing the file). After unpacking it, it is downsized for presentation purposes with the following terminal command:

```
split -l 2181785 -d taxi_february.txt taxidataS
```

As we are only interested in the first file, the redundant split files are removed:

```
for i in {01..10}; do rm taxidataS$i; done
```

We would like to split the reduced file into a main part (2.1 million data points) and a small rest (80.000 data points) for later updates. This is done with a split command:

```
split -l 2100000 -d taxidataS00 taxidata
```

Finally, delete `taxidataS00` and move the files `taxidata00` and `taxidata01` into a newly created directory `secondo/bin/cabsrome`.