Workflows for Simulating Meteorological Phenomena using the WRF Model

Ladislav Hluchý, Viera Šipková
Institute of Informatics, Slovak Academy of Sciences, Bratislava
Andrej Lúčny, Martin Gažák
Microstep-MIS, Monitoring and Information Systems, Bratislava

Slovak Research and Development Agency

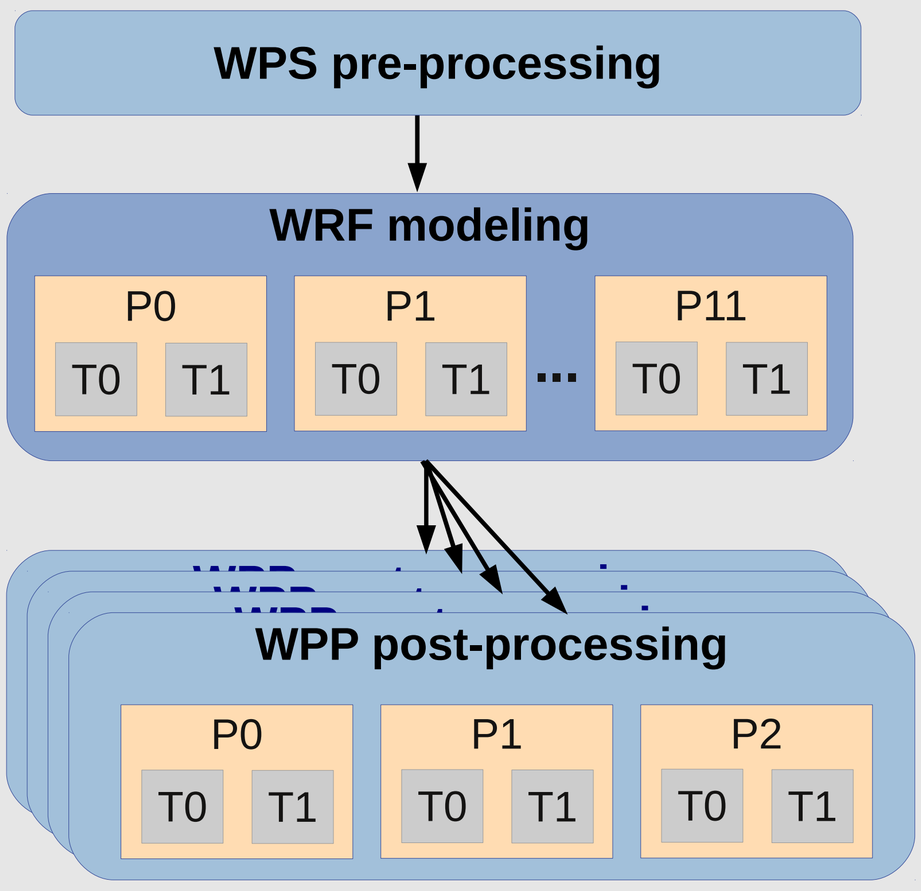
Project DMM (Data Mining Meteo)

- **Funded:** SRDA (2009-2011)
- **Partners:** Microstep-MIS, II SAS, TUKE
- **Objective:** Short-term weather warnings
- **Schema:**
  - Research of parameterized prediction models and methods for detection of significant meteorological events (rain, fog, low cloud cover)
  - Detection and prediction are based on statistical and climatological methods combined with knowledge discovery (data mining)
- **WRF model:**
  - Used to generate input data for training the data mining model

WRF Workflow Definition

A sequence of three consecutive jobs, where the input and execution of the second/third job is dependent on the output of the first/second job

- **Representation**: DAG (Directed Acyclic Graph)
- **Individual jobs:**
  - WPS pre-processing: normal job comprising of a set of serial programs
  - WRF modeling: parallel job comprising of two parallel programs (initialization and integration)
  - WPP post-processing: parallel job comprising of one parallel program running in a cycle (parametric study)



- **Technologies:** GNU compilers, C-Shell, PBS batch system, gLite middleware

Weather Research and Forecasting

A mesoscale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs

Key components:

- **Pre-processing system (WPS)** – its function is to transform large-scale terrestrial and meteorological data into a format suitable for ingest by the ARW's real data pre-processor.
- **Advanced Research WRF (ARW)** – dynamic Eulerian mass solver, the core of the modeling system
- **Post-processing (WPP)** and **visualization tools** – their function is to post-process and display the ARW simulation output data

The WRF model has been developed to run on a variety of platforms, either serially or in parallel (MPI), with or without multi-threading (OpenMP).

Running times on 2-core cluster

4 domains: Bratislava, Slovakia, Czech-Slovakia+neighbourhood, Europe
Prediction time: 72 hours

Job	Platform	#CPUs	#MPI procs	#OpenMP threads	Time hh:mm:ss
WPS	Seq	1	x	x	00:03:34
(1) WRF	DM	12	12	0	01:34:07
(2) WRF	DM+SM	12	6	2	01:16:18
(3) WRF	DM+SM	24	12	2	01:06:09
WPP 4 jobs	DM+SM	6	3	2	00:03:48
Total	DM	12	12	0	01:42:05
Total	DM+SM	24	12	2	01:13:31

Grid Overhead times

Workflow	Total	Execution	Data transfer	Grid Overhead
1 job	02:09:06	02:02:31	00:04:03	00:06:35
DAG	02:46:47	02:22:49	00:10:20	00:23:59
3 jobs	02:53:07	02:25:47	00:14:58	00:27:20

WRF Workflow Execution

- Workflow-manager shell-script – used to automate the process of workflow runs:
  - It accepts input arguments
  - It produces job-submission scripts
  - It provides for the execution of the workflow through the invocation of the job-submission scripts

This work is supported by projects: EGEE-III FP7-222667, EGI-InSPIRE RI-261323, VEGA 2/0211/09, and SRDA VMSP-P-0048-09