

# Parallel Virtual Machines in Kepler

Daniel Zinn, Xuan Li, Bertram Ludaescher

## Abstract

Scientific Workflow Systems provide an intuitive and easy interface for users to design applications for analysis and processing of data. They greatly simplify the understanding and sharing of complex computation processes and do not require the user to have any prior programming experience. An often-used model of computation for Scientific Workflows is "Process Networks" (PN), in which concurrently running processes communicate with each other via pre-defined channels. We developed a light-weight parallel PN engine, called PWS, that allows to execute PN workflows in a cluster environment. We furthermore extended Kepler to function as a front-end to PWS. The users are thus able to use the intuitive Kepler GUI (an adoption of Vergil, the Ptolemy II GUI) to build and execute PN workflows that can run locally as well as in a cluster environment utilizing parallel resources. Our demonstration will show how a movie transformation pipeline can be built within Kepler and then be executed locally as well as on a cluster computer. Kepler is loosely coupled to the cluster PN and thus the data processed on the cluster is not routed through Kepler. Instead, Kepler is used only to setup, launch and monitor cluster PN workflows. However, by exploiting the communication channel from Kepler to PWS, Kepler actors and PWS actors, coexisting in the same workflow, can communicate with each other through String ports.