Scientific Workflow Automation and Management Platform (SWAMP)

Scientific workflows have drawn an enormous amount of attention in the e-science research and development communities and be recognized as an important unifying mechanism to combine scientific data management, processing, monitoring, and visualization tasks, which are found wide acceptance in the fields of bioinformatics and cheminformatics in the early 2000s, where they successfully met the need for multiple interconnected tools, handling of multiple data formats and large data quantities. We design and develop a generic Scientific Workflow Automation and Management Platform (SWAMP), which contains a set of API libraries and easy-to-use computing and networking toolkits for application scientists to conveniently assemble, execute, monitor, and control complex computing workflows in heterogeneous high-performance network environments. SWAMP integrates the graphical user interface of Kepler to composite workflows and DAGMan for workflow dispatching and execution. SWAMP includes a web interface to configure, dispatch, monitor and debug the workflows, and a workflow mapper to map the abstract workflows real networks to optimize the workflow performance. A real case study is presented to show the feasibility of the platform.