

Towards complex and intelligent workflow programming for Distributed Computing



This research line aims to ease the development and execution of applications for distributed infrastructures, such as Clusters, Grids and Clouds.

Summary

The aim of this research line is developing programming models to ease the development of applications for distributed infrastructures, and runtime systems to execute efficiently applications in different distributed platforms in a transparent way and exploiting the inherent parallelism of the application.

For the sake of programming productivity we focus in the following key characteristics:

- **Sequential programming:** Programmers do not need to deal with the typical duties of parallelization and distribution, such as thread creation and synchronization, data distribution, messaging or fault tolerance. Instead, the model is based on sequential programming, which makes it appealing to users that either lack parallel programming expertise or are looking for better programmability.
- **Infrastructure agnostic:** The programming model abstracts the application from the underlying distributed infrastructure. Hence, Programmers do not need to include any detail that could tie them to a particular platform, like deployment or resource management. This makes applications portable between infrastructures with diverse characteristics.
- **Standard programming languages:** Programmers can use existing programming languages like Java, Python and C/C++. This facilitates the learning of the model, since programmers can reuse most of their previous knowledge.
- **APIs:** Programmers do not require to use any special API call, pragma or construct in the application; everything is pure standard Java syntax and libraries. With regard the Python and C/C++ bindings, a small set of API calls should be used on the COMPSS applications.

Objectives

- Do research and development on programming tools to **facilitate the implementation of distributed applications** in order to increase the developer productivity.
- Do research and development on runtime systems for **efficient application execution** on distributed computing environments, including the following topics:
 - * **Automatic application parallelization**
 - Do research in methodologies to detect the application inherent parallelism by detecting data dependencies between different part of code
 - * **Runtime scheduling**
 - Investigate different scheduling policies (energy-aware, data locality,...) to achieve an efficient execution of applications
 - * **Middleware Interoperability: Program once, execute everywhere**
 - Develop mechanisms to execute applications in different distributed platforms in a transparent way. In other words, The programmer do not need to make any change in the application code to execute to port the application to different execution environments.

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

Source URL (retrieved on 23 Jul 2024 - 15:00): <https://www.bsc.es/es/research-development/research-areas/distributed-computing/towards-complex-and-intelligent-workflow>