Advanced data management for geographically distributed scientific workflows on clouds

The KerData team is focusing on scalable Big Data storage and processing on clouds and post-Petascale platforms, according to the current needs and requirements of data-intensive applications [1]. In particular, this PhD proposal aims to devise a solution for the scalable management of Big Data in the context of scientific workflows executed on geographically distributed cloud datacenters or on federated and hybrid clouds. This work will be done in collaboration with researchers in the fields of workflow modeling, specification and scheduling and with several cloud providers, in particular Microsoft.

Scientific applications are often represented by workflows, which describe sequences of tasks (computations) and data dependencies between these tasks. Several scientific workflow environments have been already proposed. They have been developed primarily to simplify the design and the execution of a set of tasks in parallel-distributed infrastructures. These environments propose a "process-oriented" design approach, where the information about data dependencies (data flow) is purely syntactic. In addition, the targeted execution infrastructures are mostly computation-oriented, like clusters and grids. They have no or little support for efficiently managing large data sets. Finally, data analyzed and produced by a scientific workflow are often stored in loosely structured files. Simple and classical mechanisms for their management are used: they are either stored on a centralized disk or directly transferred between tasks. This approach is not suitable for data-centric applications (because of inherent bottlenecks, costly data transfers, etc.).

Clouds have recently emerged as an interesting infrastructure option for deploying scientific workflows. Clouds are typically composed of several geographically distributed datacenters, connected through high speed networks. Having these high-speed networks in place between datacenters is important but not sufficient. Being able to effectively use them has become increasingly important for wide-area data replication as well as for federated cloud computing. Building on the clouds elasticity, in recent years, scientific workflows have become an archetype to model experiments on such infrastructures. In addition, the Cloud allows users to simply outsource data storage and application executions. Still, there are substantial challenges ahead that must be addressed before their potential can be fully exploited in the context of scientific workflows.

One missing link is data management, as clouds mainly target web and business applications, and lack specific support for data-intensive scientific workflows. Currently, the management of workflow data in the clouds is achieved using either some application specific overlays that map the output of one task to the input of another in a pipeline fashion, or, more recently, by leveraging the MapReduce programming model (e.g. Amazon Elastic MapReduce, Hadoop on Azure - HDInsight). However, most scientific applications do not fit this model and require a more general data and task orchestration model, independent of any programming model.

The goal of this PhD thesis is to specifically address these challenges by proposing a framework for the efficient data management of data-intensive workflows in clouds and federated clouds. We observed that these scientific workflows have become increasingly complex and more demanding in terms of their computational and data requirements, some generating data volumes exceeding petabytes; clearly they can no longer be executed within a single datacenter – cloud providers tend to distribute intensive workloads in order to avoid resource contention and minimize the impact of multi-tenancy between users. In this context, these workflows are executed on several geographically distributed datacenters concurrently or even on several federated clouds. Since there is no support for the efficient handling of data across several cloud datacenters, this PhD proposal aims to fill this gap.
Our approach will leverage the cloud infrastructure capabilities for handling and processing large data volumes. In order to support data-intensive workflows, our cloud-based solution will: 1) Adapt the workflows to the cloud environment and exploit its capabilities; 2) Optimize data transfers to provide reasonable times; 3) Manage data and tasks so that they can be efficiently placed and accessed during execution.

This thesis will focus on data storage and processing issues, in particular leveraging data and workflow awareness for efficient sharing and accessing large dataset under high concurrency. This approach will rely on collocating data on the compute nodes, in order to exploit data locality and to avoid the overhead of interacting with a shared file system. We have already proposed an approach for federating the virtual disks into a scalable globally shared object store and validated it with MapReduce [2]. The next step consists in devising a solution for more general workflows, independent of a specific programming model. This proposal will rely on the fact that the scientific workflows generate highly similar access patterns to the data. Under these circumstances, the targeted solution will enable high throughput data transfers without requiring any foreknowledge of the access pattern and will enable an autonomic behavior of the cloud-based workflow based by dynamically adapting to the workflow context (data size, format, access, resource cost).

Furthermore, the PhD thesis will mitigate the large-scale end-to-end data movement bottleneck over wide-area by efficiently utilizing underlying networks and effectively scheduling and optimizing data transfer tasks. To this end, we will design a cloud data-infrastructure model. This implies the study of how to compute a cost for possible inter and intra sites (clouds) communication graphs, storage and computation resources for a workflow on given cloud infrastructure. Data storage and transfer strategies that will be used to compute an abstract execution plan should minimize the cost (with eventually a maximum execution time).

The validation of this proposal will be performed using synthetic benchmarks and real-life applications from bioinformatics: first on the Grid5000 platform in a preliminary phase, then on the Microsoft Azure cloud environment, in collaboration with the Microsoft Azure team. The goal to push the Microsoft data centers to their limits in order to run large-scale scientific workflows, store huge volumes of data, and share the results with the scientific community.

**Dpt scientifique** • Scientific department: D1 - Systèmes large Échelle / Large Scale Systems

**Équipe projet** • Research team: KerData ; [http://www.irisa.fr/kerdata/doku.php?id=open_positions](http://www.irisa.fr/kerdata/doku.php?id=open_positions)

**Directeur de thèse** • PhD Director: Gabriel Antoniu

**Encadrant(s)** • PhD supervisors: Gabriel Antoniu / Alexandru Costan

**Contact(s)** : gabriel.antoniu@inria.fr / alexandru.costan@irisa.fr

**Début des travaux** • Work start date: 01/10/2013

**Lieu** • Place: IRISA - Rennes
References:
