# ARC 6 : Technologies de l'Information et de la Communication et Usages Informatiques Innovants

Trusted SLA-Guided Data Integration on Multi-cloud Environment

The emergence of cloud environments redefined the data integration as a service matching and composition problem. Furthermore, new constraints imposed by multi-cloud environments bring new challenges to data integration

Objective: propose a data integration approach in a multi-cloud environment taking into account data quality properties and service level agreements (SLA)
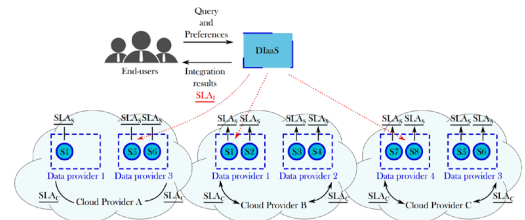
POINTS CLES

- Model quality measures linked to data and to cloud resources
- SLA for guiding the data integration process performed on a multi-cloud
- Match data providers (services) to queries for integrating data according to the user requirements and the SLAs

PROBLEMATIQUE

Efficiently obtain results for queries addressing services deployed in a multi-cloud environment by: *(a)* Fulfilling user QoS requirements; *(b)* Respecting subscribed contracts with the involved cloud provider(s); and *(c)* Respecting services contracts (constraints).
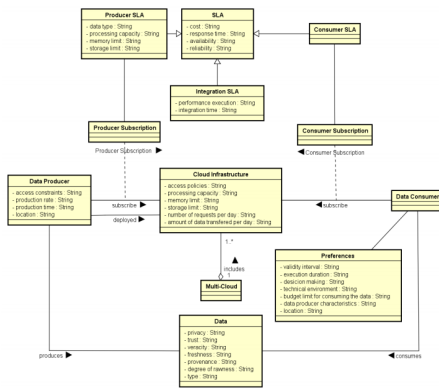

Figure 1: Data integration metamodel

Challenges:
- How to be sure that all the agreed SLAs are respected while satisfying the user?
- Can my constraints be satisfied? Which services shall I ask for?
- How can resources be saved for next query?
- How to perform the query rewriting matching services that answer the query and satisfy the quality preferences?
- How to integrate different SLA associated to services involved with user's quality preferences?

TRAVAUX ENGAGES / RESULTATS

Problem statement and state of the art by applying a systematic mapping process

Rhône algorithm: formalization and implementation of a quality-based query rewriting considering user preferences and quality aspects expressed as SLA contracts (Figure 2)
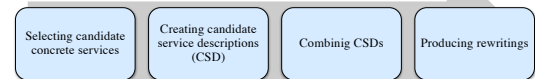

Figure 2: Rhone algorithm phases

Experimentation: lessons learned (Figure 3)

Complexity:
- Selecting candidate concrete services: O (n^2)
- Creating candidate service descriptions: O (n^3)
- Combining CSDs: O $(n^k)^m$

Performance increased reducing
- The number of rewriting solutions
- Integration execution time
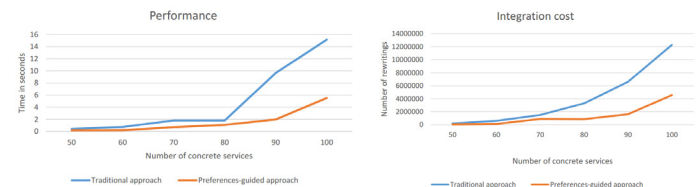
Rewritings quality enhanced

Figure 3: Evaluation results

Data integration metamodel: a metamodel and metaprocess to address data integration adapted to the multi-cloud context (Figure 1)

Ongoing work:

Designing models for the *cloud SLA*, *service SLA* and *integration SLA*

Reducing the overhead caused by query rewriting:
- Taxonomy of query variations for promoting **reusability** of rewriting results
- **Heuristics for optimizing** the rewriting approach adapted to the multi-cloud context

Evaluation of the overall data integration approach adapted to the multi-cloud context

Publications

Daniel Aguiar da Silva Carvalho, Magellan, IAE, Univ. J-Moulin Lyon 3 - France
Chirine Ghedira Guegan, Magellan, IAE, Univ. J-Moulin Lyon 3 - France
Nadia Bennani, CNRS INSA-Lyon, LIRIS, UMR 5205 - France
Genoveva Vargas-Solar, CNRS, LIG (UMR 5217) - LAFMIA (UMI 3175) - France