Integrating databases on a multi-cloud environment using SLA contracts
SLA-based secure database integration on a multi-cloud environment
Through secure database integration using SLA in multi-cloud context

Nadia Bennani , Chirine Ghedira-Guegan[y], Genoveva Vargas-Solar[z] and Martin A. Musicante[x]

Univ. Lyon, CNRS INSA-Lyon, LIRIS, UMR5205, France *(nadia.bennani@insa-lyon.fr)*
[y]MAGELLAN, IAE, Univ. J-Moulin, Lyon 3, France *(chirine.ghedira-guegan@univ-lyon3.fr)*
[z]CNRS, LIG-LAFMIA, St. Martin D'Hères, France *(genoveva.vargas@imag.fr)*
[x]DIMAp, UFRN, Natal, Brazil *(mam@dimap.ufrn.br)*

Abstract—**The recent emergence of the cloud paradigm opens new challenges for data processing. Indeed, unlimited access to cloud resources and the "pay as U go" model change the hypothesis for processing big data collections. This is accentuated due to cloud provider incentives to attract data providers and let them choose their cloud. In this context data services are deployed on several clouds which make data integration face several heterogeneity problems. This paper proposes a secure-aware data integration (lookup, aggregation, correlation) strategies that can be performed on a multi-cloud environment according to their respective SLA contracts. The resulting data integration is provided as a DaaS (data as a service) with a global SLA that consolidates as much as possible the SLA adopted by every cloud provider. Our data integration approach has been tested for integrating databases used for managing an education program.**

Keywords—*SLA; Cloud Computing; Data Integration;*

I. INTRODUCTION

This paper presents an approach for addressing data collections integration in a multi-cloud hybrid context guided by user preferences statements and SLA contracts exported by different cloud providers. The objective is to propose an SLA guided data provision and integration system exported as a DaaS by a cloud provider that can be adapted according to user preferences on the way data can be delivered such as accepting partial results delivered on demand or under predefined subscription models that can affect the quality of the results; accepting specific data duplication that can respect privacy but ensure data availability; accepting to launch a task that contributes to an integration on a first cloud whose SLA verifies a QoS requirement rather that a more powerful one.

Integrating and processing heterogeneous data collections, calls for efficient methods for correlating, associating, filtering them taking into consideration their "structural" characteristics (due to the different data models) but also their quality, e.g., trust, freshness, provenance, partial or total consistency. Existing data integration techniques have to be revisited considering weakly curated and modeled data sets. This can be done according to quality of service requirements expressed by their consumers and Service Level Agreement (SLA) contracts exported by the cloud providers that host these collections and deliver resources for executing the associated management processes.

In this context, this paper presents an approach proposing strategies for computing integrated SLA's according to agreed SLA's exported by services and adaptable query rewriting for integrating data sets according to user preference statements. This implies to consider several granularities of SLA: first, at the cloud level; the SLA ensured by providers regarding data; then at the service level, as unit for accessing and processing data, to be sure to fit particular service needs; and finally at the integration level i.e the possibility to process, correlate and integrate big data collections distributed across different cloud storage supports, providing different quality properties to data (trust, privacy, reliability, etc).

Accordingly, the remainder of this paper is organized as follows. Section II presents related works that address SLA modelling, integration and SLA guided data management processes. Section III gives an overview of our approach for integrating data sets provided by services (i.e., DaaS) by concilating SLA's provided by services and user's profiles expressing QoS preferences about the data they want to consume and the conditions in which they must be processed and delivered. Section VII presents a use case for illustrating the interest and use of our approach. Finally VIII concludes the papers and discusses future work.

II. RELATED WORK

The advent of cloud computing imposed a new model for resource consuming that is not concerned with resource