2011 First International Symposium on Network Cloud Computing and Applications

A Unified Management Model for Data Intensive Storage Clouds

Athanasios Voulodimos[1], Spyridon V. Gogouvitis[1], Nikoletta Mavrogeorgi[1], Roman Talyansky[2], Dimosthenis Kyriazis[1], Stefanos Koutsoutos[1], Vasileios Alexandrou[1], Elliot Kolodner[3], Per Brand[4], Theodora Varvarigou[1]
[1]School of Electrical and Computer Engineering, National Technical University of Athens
Athens, Greece
[2]SAP Research Israel
Ra'anana, Israel
[3]IBM Haifa Research Lab
Haifa, Israel
[4]Swedish Institute of Computer Science
Kista, Sweden
E-mail addresses: {thanosv, spyrosg, nikimav}@mail.ntua.gr, roman.talyansky@sap.com, {dimos, stefk, alexv}@mail.ntua.gr, kolodner@il.ibm.com, perbrand@sics.se, dora@telecom.ntua.gr

***Abstract*—Cloud storage has recently appeared as a promising solution to handle the immense volume of data produced in nowadays' rich and ubiquitous digital environment. Existing approaches, however, do not fully address certain significant issues, such as Service Level Agreements, which tend to be rudimentary, and management of cloud systems, which often lacks sophistication. In this paper, we propose the definition and development of management models that systematically describe the resources, services, usage, and requirements in the context of a cloud storage environment. Furthermore, we introduce content oriented SLA Schemas, which are adapted according to the content term involved in each SLA. Combining all this often heterogeneous information in a smart, structured, and uniform manner can provide significant added value, allowing correlations, links and extraction of useful inferences. We thus propose to incorporate these models in a rich Unified Management Model, which will contribute to optimum decision making, data movement, as well as better resource allocation and energy management, also leading to enhanced Quality of Service. The research presented in this paper is work in progress carried out in the context of the VISION Cloud project.**

I. Introduction

The growth of digital data during the past years has increased in a such rate that is challenging our ability of storing it. Storage Cloud solutions, by which storage is virtualized and offered on demand, promise to address the proliferation of data. Aspects of Cloud computing, such as the pay-as-you-go model, essentially unlimited scalability and capacity, lower costs are also prominent in Storage Clouds. Especially important to the uptake of such solutions is their ease of use and management. In this context VISION Cloud [1] has presented the architecture of a scalable and flexible cloud environment addressing the challenge of providing data-intensive storage cloud services through raising the abstraction level of storage, enabling data mobility across providers, allowing computational and content-centric access to storage and deploying new data-oriented mechanisms for QoS and security guarantees [2].

The QoS mechanisms and SLAs offered by current cloud storage providers are rather rudimentary. They guarantee only service availability and give service credit or refunds for lack thereof [3], [4], [5], but do not address data availability and protection. In research, an architecture and protocol for an SLA-based trust model for cloud computing have been presented in [6] and approaches for managing

978-0-7695-4550-9/11 $26.00 © 2011 IEEE
DOI 10.1109/NCCA.2011.18

the mappings of low-level resource metrics to high-level SLAs in [7], [8], [9] have been proposed.

In our approach, models, requirements and SLA Schemas are expressed not only on storage resources and services, but also on the content descriptions for the underlying storage objects, in support of content centric storage. Moreover, models in VISION Cloud will combine resource and data content information providing a cross-discipline combination both for storage resources and data content characteristics in order to be used as an integrated approach.

Therefore the desired functionality with the guaranteed QoS levels can be implemented, while assuring that service utilization is improved. We focus on modeling the resources by aggregating information related to them (e.g. expressing capabilities such as read / write performance) with information related to the content (e.g. content type, request frequency, etc) as well as capturing information regarding usage and requirements parameters with the final goal of producing a Unified Model, able to capture all information related to management. The aforementioned unified model allows for optimum management of resources and services given their rich representation and characterization. For example, in the cases of frequent access / request to a specific service the infrastructure will be adopted to meet the demand by increasing the bandwidth to the resource. The models produced also express the importance of a service and a resource in comparison to other ones and can consist as input for management decisions, such as geographic placement of data or security mechanisms that shall be put in place to protect the specific "high importance" resources and services etc.

II. Resource Modeling

Modeling the resources in VISION Cloud contributes to the efficient management of nodes within a cluster, of clusters within a data center, and of data centers within a cloud. It offers a standardized, organized way to have an overview of all resources within the cloud.

A set of parameters for each level of the physical model (e.g. node, cluster, cloud) is defined. The modeling of nodes within the platform is consisted of parameters related to the identity of the node, its storage, compute and network capabilities. The same format of distinguishing between identity, storage, compute and network parameters is also used in modeling a cluster. A data center comprises of one or, usually, more clusters. Since the emphasis in VISION Cloud is more on clusters than on data centers, the data center model is an aggregation information regarding the clusters included with the addition of energy consumption information.

III. Service Modeling

Modeling of services deployed and provided within VISION Cloud also constitute part of the management models. Presently within the scope of our platform we can differentiate between two main categories of services: data object uploading/ingestion, access and update, and storlets. A storlet is a construct which specifies a computation to be executed on a set of data objects. The two most important characteristics of storlets are that they execute close to their data and that the style of computation is not limited to batch job executions.

For the upload/access/update services, significant parameters include the resource requirements needed, the content type the service refers to, the permission of the data object, the device used to access a data object and key performance incicators used to evaluate the success of service provisioning.

Regarding storlets, the elements that are of significance are determined in the computational storage model and include information such as a description of the computation to be executed, the idenity of the associated data objects, constraints on the storlet (such as maximum working memory), triggers defining when and if a strolet is activated and credentials that the storlet should use to access data objects.

IV. Usage Modeling

Usage characteristics (e.g. geographical access distribution or read / write access frequency) encode the knowledge that is automatically harvested and captures application access patterns with regard to storage. These characteristics are of major importance since they may be used afterwards for optimising mapping of cloud storage resources to application requirements. In addition to automatic harvesting of usage characteristics, applications may inject into the cloud storage infrastructure a priori knowledge on the application behaviour. Self knowledge may be used to further optimise mapping of application requirements to cloud storage resources. Therefore, usage models include representations and descriptions to inject application self-knowledge. More specifically, the model is used to build predictive models for the Planner part of the MAPE loop [10].

To build a usage model a time axis is split into time intervals of the same duration, e.g. hours, days or weeks. For each interval a vector of attributes is calculated from the aggregated monitoring information of this interval. The attributes describe the application's characteristics such as total the number of reads and writes, the size and standard deviation of accessed objects, etc. Thus a vector of attribute values is generated for each time interval. A usage model consists of all the attribute vectors that were generated in the past.

V. Requirements Modeling

An application running over the VISION Cloud infrastructure will be able to specify requirements, which will be later used by the infrastructure to drive the data access operations of the application. To this direction, a Requirements Model capturing both the requirements emerging from application attributes modeling and the ones included in the SLAs is used.

Applications specify attributes that are automatically translated to storage requirements. For example, attributes such as "Master Data" or "Transactional Data" are translated into requirements "low update rates, replicated with eventual consistency" and "high update rates, strong consistency" respectively. Once the set of requirements is identified, the methods to specify their values need to be provided. Such methods comprise generic models for storage requirements. For example, high availability may be specified by the number of replicas required, as well as the coarse geographical location for each replica. It should be noted that while storage requirements are technical in nature and they reflect the applications' requirements, SLA Schemas have a more business-oriented flavour (e.g., as they specify penalties in case of SLA breaches). However, given that the SLA Schemas will include content-related terms, these will be considered as additional storage requirements.

VI. SLA Schemas

An SLA is a service level agreement between the tenant and the service provider, where a contract is created that defines the terms and conditions under which the services will be provided. In VISION Cloud, the SLAs are enriched versions of conventional ones. They contain not only requirements related to cost, availability, performance, etc., but also content related terms, in the context of VISION Cloud's content centric approach. For example, the pricing models and the performance estimation for data transfer is different depending on the SLA content term.

An SLA Schema is a schema that represents an SLA. It defines what data can be contained in an SLA (e.g. cost, contract dates, user requirements, fines, etc) and in which format. The language that is used for the SLA Schema is XSD (XML Schema Document).

The basic SLA Schema's elements include a term attribute that represents the content term related to the SLA (e.g. media, healthcare, etc.), attributes such as availability, jitter, ingest rate, etc., cost and penalties.

VII. Unified Model

The aforementioned models and schemata contain and organize various pieces of information which are of importance to the Cloud in different ways. Combining all this often heterogeneous information in a smart, structured, and uniform manner can provide significant added value, allowing associations, links and extraction of useful inferences. We thus propose the design of a Unified Model that unifies the earlier models presented. In particular, in VISION Cloud we aim at combining all types of metadata and information characterizing different parts / "elements" of the system (data objects, resources, services, requirements) into a Unified Model, which will allow for optimum decision making, data movement, as well as resource and service management, also leading to enhanced Quality of Service.

Through the platform's Management Interface, the users (both customers - in the case of requirements / SLAs, and developers - in the case of content descriptions for the storage objects) are able to specify the different parameters needed to produce each individual model in a secure environment. These, along with parameters generated by the platform itself (such as usage parameters) are automatically aggregated to form the Unified Model.

The Unified Model's usefulness lies in the fact that various kinds of information pertaining to different aspects of interest to VISION Cloud are organized, structured, and presented in a standardized, formalized, consistent, and unified manner. It enables the ability to extract correlations and make inferences, linking information stored in different models. For example, the significance of a resource can be determined not only from its "obvious" hardware characteristics (Resource Model), but also based on the content stored in that resource, the usage patterns linked to it (Usage Model), the triggers or constraints of the storlets that could possibly be associated with this content (Service Model), or even the strictness of the SLA terms connected to the data.

Moreover, one of the goals of VISION Cloud is to substantially extend the limited data migration capabilities available in contemporary infrastructures so as to migrate and federate data across geographically distributed administrative domains, thus ensuring comprehensive and transparent data interoperability and overcoming the problem of data lock-in. A Unified Management Model helps in federation scenarios by providing a common way in which all aspects of a platform can be defined. In this way, knowledge that is accumulated over long periods of time regarding particular data objects hosted on one provider can be easily migrated to a new provider.

VIII. Conclusion

We have presented a series of management models aimed at capturing the most crucial information in the context of a cloud environment for data intensive storage services. Through the modeling of resources, services, usage and requirements, together with the definition of content oriented SLAs, we aim at providing a standardized, structured and uniform way to organize the important information involved in the cloud. Through a less conventional, more "modern" notion of metadata model, which is inclusive and emphasizes content, our goal is to produce rich, useful models, and smartly combine them in a Unified Management Model, which will be able to express the importance of data and/or resources in comparison to other data and/or resources and to efficiently provide input to decision making and placement mechanisms. This contributes to better resource allocation and energy management within the platform. Currently the VISION Cloud project is in its early stages, thus the implementation and evaluation of the concepts outlined in this paper are underway. Discussing outcomes and results is deferred for future work.

Acknowledgment

The research leading to these results is partially supported by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 257019, in the context of the VISION Cloud Project.

References

[1] "VISION Cloud project." [Online]. Available: http://www.visioncloud.eu/

[2] H. Kolodner, D. Naor, S. Tal, S. Koutsoutos, N. Mavrogeorgi, S. Gogouvitis, D. Kyriazis, and E. Salant, "Data-intensive Storage Services on Clouds: Limitations, Challenges and Enablers," in *eChallenges e-2011 Conference*, 2011.

[3] "Amazon S3 service level agreement," 2009. [Online]. Available: http://aws.amazon.com/s3-sla/.

[4] "Nirvanix Service Level Agreement." [Online]. Available: http://www.nirvanix.com/sla.aspx

[5] M. Corporation, "Windows Azure Pricing and Service Agreement," 2009. [Online]. Available: http://www.microsoft.com/windowsazure/pricing/

[6] M. Alhamad, T. Dillon, and E. Chang, "Conceptual SLA framework for cloud computing," in *Digital Ecosystems and Technologies (DEST), 2010 4th IEEE International Conference on*, april 2010, pp. 606 –610.

[7] I. Brandic, V. C. Emeakaroha, M. Maurer, S. Dustdar, S. Acs, A. Kertesz, and G. Kecskemeti, "LAYSI: A Layered Approach for SLA-Violation Propagation in Self-Manageable Cloud Infrastructures," *Computer Software and Applications Conference Workshops*, vol. 0, pp. 365–370, 2010.

[8] V. C. Emeakaroha, I. Br, M. Maurer, and S. Dustdar, "Low Level Metrics to High Level SLAs - LoM2HiS framework: Bridging the gap between monitored metrics and," in *SLA parameters in Cloud environments. The 2010 High Performance Computing and Simulation Conference (HPCS 2010) June 28July 2, 2010.*

[9] S. Gogouvitis, K. Konstanteli, S. Waldschmidt, G. Kousiouris, G. Katsaros, A. Menychtas, D. Kyriazis, and T. Varvarigou, "Workflow management for soft real-time interactive applications in virtualized environments," *Future Generation Computer Systems*, vol. 28, no. 1, pp. 193 – 209, 2012.

[10] IBM, "An architectural blueprint for autonomic computing." April 2003.