Orsay, April 2011

Efficient cloud-based stores for large-scale Web data management

Leo joint team, INRIA Saclay
Supervisors: Ioana Manolescu, Dario Colazzo and Nicole Bidoit

The recent development of commercial and scientific cloud computing environments has strongly impacted research and development in distributed software platforms. From a research perspective, the main interest of the cloud is the opportunity for massive parallelism in storing and manipulating data, and in running complex computations. From a business perspective, an extra advantage of cloud platforms is freeing the tenant of the application (e.g., whose computation or Web presence is hosted in the cloud) from the burden of administering the hardware: cloud platforms guarantee quasi-unbounded resource allocation, elastic scaling up and down of the resources according to the demand, and reliable execution (tasks started in the cloud will complete even after a possible failure). In parallel with the development of cloud hardware platforms, software infrastructures for massive parallelism have been developing since the original MapReduce model [DG04].
Cloud infrastructures and MapReduce-style parallel programming have received attention in many CS research areas, see, e.g., [SoCC2010]. In the data management field, the focus has been on building "MapReduce-plus" software frameworks providing higher-level abstractions (tuples, data-processing operators…), subsequently compiled into efficient MapReduce programs [AVH+11, VCG+11]. However, these efforts focus on the lower levels of a data processing engine, without reaching to the higher level of expressive database query and update languages such as XQuery (for XML) or SPARQL (for the Semantic Web data language RDF). Commercial database vendors have ported their existing products to cloud infrastructures, by scaling out the lower layers of the data store [MS2011]. However, these engines mostly focus on relational data, and do not cater well to XML and RDF databases of Web data.

Cloud-related work in Leo and ICT Labs: The Leo team focused on expressive models and efficient platforms for distributed Web data management, and in particular (the potential supervisor of this thesis) on efficient XML processing. A particular success in distributed Web data management is materialized by the ViP2P platform, allowing efficient and scalable sharing of XML and RDF data. ViP2P incorporates advanced database indexing and query rewriting techniques, and has been shown to scale very well on Terrabytes of XML data and hundreds of peers.
Since 2010, Leo has started to study the issues raised by the efficient management of XML and RDF data in a cloud-based distributed environment. This research involves three permanent staff (I. Manolescu, D. Colazzo and N. Bidoit) and two interns. An ADT ("*Action de Développement Technologique*", junior engineer INRIA position) has been secured for 2011-2013 to organize and strengthen our code development in this area.
Leo is a participant to a "Connected Clouds in the City" 2011 activity of the KIC EIT ICT Labs, within the Thematic Action Line "Digital Cities". Moreover, we will participate to (and probably coordinate) a proposed 2012 activity called "Data-intensive Cloud Infrastructure" within the "Cloud Computing" ICT Labs action line. This submission has been discussed and is encouraged by the action line coordinator (Seif Haridi). It complements other planned activity proposals for "Cloud

Computing", related to topics such as reliability, benchmarking, cloud data usage mining etc.) Collaborators in the "Data-intensive Cloud Infrastructure" activity are Volker Markl and his group from the Technical University of Berlin, as well as the team around Alexandru Iosup from the Technical University of Delft.

Proposed work: The PhD topic we propose concerns *devising algorithms and a platform for efficient cloud-based Web data stores*, in particular focusing on XML (structured documents) and RDF (semantics and knowledge) data. It has been by now widely accepted that a huge fraction of the world's interesting data is either shared through the Web (data-rich Web sites such as organization and product catalogs…) or directly produced through and for Web platforms (social media sites, RSS and tweets, blogs, online publications,…) Many organizations now recognize the value of the trove of Web data and the need for scalable platforms to manage it. [1]
We plan to research and build a *single check-point data management application, where users will be able to upload, query, update, and reason about the data all the while it is remotely stored and managed in a cloud platform*, taking advantage of its elasticity and reliability features. This task raises two main challenges:

- Devising a **multi-layered data storage and indexing model**, taking into account the existing services provided by cloud platforms: Cloud-based stores provide multiple services recommended for distinct usage patterns. For instance, Amazon proposes S3 as a general storage platform for voluminous objects, with relatively high response time, as well as SimpleDB for storing numerous small data objects with faster access but several stringent size limitations. Our data storage architecture should adapt to such different functionalities in a customizable and generic way, to enable its deployment on varying and possibly distinct heterogeneous cloud platforms.
- Conceiving **efficient algorithms for processing queries and updates** on Web data: in this area, we will both investigate building specific algorithms for our needs relying directly on the cloud service APIs, and leveraging the platform of our ICT Labs activity partner TU Berlin (Stratosphere) in order to take advantage of the efficient parallel execution primitives (building on MapReduce) which they have built. Particular attention will be given to meaningful models for performance and monetary costs associated to cloud-based hosting and processing of large data. Studies on cost policy-aware replication [DKA09] and on cloud-based transaction performance [KKL10, IYE11] will be our starting points.

Profile and competencies: We seek a candidate with an excellent academic record and in particular strong credentials in the areas of data management and Web technologies. Of particular interest is the ability (and ideally demonstration) of ease in the principled development of a large software system.
The student will work in an international team and is expected to frequently interact with others and in particular with the ADT engineer whose efforts will back up and consolidate his research-oriented efforts; the candidate will also have to interact and/or visit our ICT Lab partners. Excellent English and communication skills are thus a must.

[1] Eric Schmidt (CEO of Google) notes : « *Every two days now, we create as much data as we did from the dawn of civilization until 2003. A lot of the new data is not locked away in enterprise databases, but is freely available to the world in the form of social media: status updates, tweets, blogs, and vidéos* »

References

[AVH+11] A. Alexandrov and S. Ewen and M. Heimel and F. Hueske and O. Kao and V. Markl and E. Nejcamp and D. Warneke. "MapReduce and Pact – comparing parallel programming models", BTW workshop, 2011. See also the Stratosphere page: http://stratosphere.eu

[DKA09] D. Dash, V. Kantere and A. Ailamaki: "An Economic Model for Self-Tuned Cloud Caching", in the IEEE ICDE conference, 2009

[DG04] J. Dean an S. Ghanavat. "MapReduce: Simplified Data Processing on Large Clusters", Sixth Symposium on Operating System Design and Implementation, 2004.

[KKL10] D. Kossmann, T. Kraska, S. Loesing: "An evaluation of alternative architectures for transaction processing in the cloud", in the ACM SIGMOD Conference, 2010

[IYE11] A. Iosup, N. Yigitbasi, and D. Epema, "On the Performance Variability of Production Cloud Services", in the IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing (CCGrid 2011)

[MS2011] P. Bernstein et al. "Adapting Microsoft SQL Server for Cloud Computing", in the IEEE ICDE 2011 conference (industrial session)

[SoCC2010] J. Hellerstein and S. Chaudhuri and M. Rosenblum, editors. Proceedings of the 1st International Symposium on Cloud Computing (in conjunction with the ACM SIGMOD conference), Indiana, USA, 2010

[VCG+11] V. Borkar and M. Carey and R. Grover and N. Onose and R. Vernica. "Hyrax: A Flexible and Extensible Foundation for Data-Intensive Computing", in the IEEE ICDE 2011 conference. See also the Hyrax page: http://sourceforge.net/projects/hyrax