Scalability in RDF Stream Processing Systems

Alejandro Llaves and Oscar Corcho

Ontology Engineering Group, Universidad Politécnica de Madrid, Madrid, Spain
{allaves,ocorcho}@fi.upm.es

There is no universal definition of *scalability* in the context of software systems [6]. In the field of RDF Stream Processing (RSP), we can consider that a system is scalable if it is able to process queries over a growing amount of input data efficiently, i.e. without losing quality of service. As a community interested in RSP, we have to (1) identify the causes and effects that affect the performance of a RSP system and (2) understand their relationship to keep the system performance as desired [4].

Our research focuses on scalable approaches to execute queries over RDF data streams. We investigate real-time processing technologies, e.g. Storm[1] or Spark[2], to parallelize the execution in distributed environments. We are also interested in self-adaptive strategies that allow systems to react to changes in the input data (e.g. frequency, number of streams, or heterogeneity), variable query complexity, and failure of processing nodes.

Dr. Oscar Corcho is an associate professor and leads the Data Integration group at the Ontology Engineering Group, Universidad Politécnica de Madrid. He has broader experience in the field of RSP, ranging from novel methods to access streams, definition of sensor observation ontologies, and RDF compression techniques [10,1,3,2,5].

Dr. Alejandro Llaves works as a postdoctoral researcher at Corcho's group. His background is on event processing in the Semantic Sensor Web [11], with publications related to semantic interoperability among information communities [7,9,8].

We are active participants of the W3C RSP community group[3] and believe that our contribution to this workshop is relevant to the ESWC audience because of the cross-domain nature of RSP, addressing fields like Distributed Systems, Social Networks, and Internet of Things.

Acknowledgements

---

[1] https://storm.apache.org/

[2] https://spark.apache.org/streaming/

[3] https://www.w3.org/community/rsp/

References

1. Calbimonte, J.P., Corcho, O., Gray, A.J.G.: Enabling ontology-based access to streaming data sources. In: Proceedings of the 9th International Semantic Web Conference on The Semantic Web - Volume Part I, pp. 96–111. Springer Berlin Heidelberg, Shanghai, China (2010)
2. Calbimonte, J.P., Jeung, H., Corcho, O., Aberer, K.: Enabling Query Technologies for the Semantic Sensor Web. International Journal on Semantic Web and Information Systems 8(1), 43–63 (2012)
3. Compton, M., Barnaghi, P., Bermudez, L., Garcia-Castro, R., Corcho, O., Cox, S., Graybeal, J., Hauswirth, M., Henson, C., Herzog, A., Huang, V., Janowicz, K., Kelsey, W.D., Phuoc, D.L., Lefort, L., Leggieri, M., Neuhaus, H., Nikolov, A., Page, K., Passant, A., Sheth, A., Taylor, K.: The ssn ontology of the w3c semantic sensor network incubator group. Web Semantics: Science, Services and Agents on the World Wide Web 17(0) (2012)
4. Duboc, L., Rosenblum, D.S., Wicks, T.: A framework for modelling and analysis of software systems scalability. In: Proceedings of the 28th International Conference on Software Engineering - ICSE '06. pp. 949–952. ACM, New York, New York, USA (2006)
5. Fernández, J.D., Llaves, A., Corcho, O.: Efficient RDF Interchange (ERI) Format for RDF Data Streams. In: Mika, P., Tudorache, T., Bernstein, A., Welty, C., Knoblock, C., Vrandei, D., Groth, P., Noy, N., Janowicz, K., Goble, C. (eds.) The Semantic Web ISWC 2014, Lecture Notes in Computer Science, vol. 8797, pp. 244–259. Springer International Publishing (2014)
6. Hill, M.D.: What is Scalability? ACM SIGARCH Computer Architecture News 18(4), 18–21 (1990)
7. Janowicz, K., Bröring, A., Stasch, C., Schade, S., Everding, T., Llaves, A.: A RESTful proxy and data model for linked sensor data. International Journal of Digital Earth pp. 1–22 (Sep 2011)
8. Llaves, A., Kuhn, W.: An Event Abstraction Layer for the Integration of Geosensor Data. International Journal of Geographical Information Science (2014)
9. Llaves, A., Michels, H., Maué, P., Roth, M.: Semantic event processing in ENVISION. In: Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics - WIMS '12. p. 1. ACM Press, New York, New York, USA (2012)
10. Sequeda, J.F., Corcho, O.: Linked Stream Data: A Position Paper. In: Proceedings of the 2nd International Workshop on Semantic Sensor Networks, SSN 09. CEUR-WS, Washington, USA (2009)
11. Sheth, A., Henson, C., Sahoo, S.S.: Semantic Sensor Web. IEEE Internet Computing 12(4), 78–83 (2008)