Report from the second workshop on Scalable Workflow Enactment Engines and Technology (SWEET'13)

Jacek Sroka
University of Warsaw, Poland
j.sroka@mimuw.edu.pl

Jan Hidders
TUDelft, The Netherlands
a.j.h.hidders@tudelft.nl

Paolo Missier
Newcastle University, UK
paolo.missier@ncl.ac.uk

ABSTRACT

This report summarizes the Second International Workshop on Scalable Workflow Enactment Engines and Technologies (SWEET'13). This workshop was held in conjunction with the 2013 SIGMOD conference in New York, NY, USA on June 23th, 2013. The goal of the workshop was to bring together researchers and practitioners to explore the state of the art in workflow-based programming for data-intensive applications, and the potential of cloud-based computing in this area. The program featured 4 paper presentations and two very well attended invited talks by Prof. Paul Watson, Newcastle University, UK and Dr Jelena Pjesivac-Grbovic from Google, Inc.

1. INTRODUCTION

The SWEET workshop is aimed at exploring the cross-over between languages and models for parallel data processing, and traditional workflow technology, primarily on a cloud infrastructure and for data-intensive applications. The next generation of these systems is increasingly capable of dealing with changing circumstances. Rather than efficiently running off-line a specific workflow on a predictable cloud-based data processing back-end, they now have to deal with dynamic behavior such as real-time data-analysis during the execution, user-interference with the computation while executing and changes in the computational efficiency or network structure of the heterogeneous execution back-end. These developments were reflected in the contributions of this edition of the workshop. For example, the first paper presents the *DynamicCloudSim* system for simulating the effects of certain resource allocation and scheduling strategies in dynamic cloud-based distributed architectures where the efficiency of the different services may change in time. The second paper presents STAFiLOS, a STreAm FLOw Scheduler, which allows the stream-based execution of a workflow with dynamic input streams on top of a conventional workflow execution engine. The third paper introduces OSIRIS-SR, a distributed peer-to-peer workflow execution framework that allows workflows to be efficiently and reliably executed on a dynamic networks of cooperating nodes. The final paper gives an overview of user-steering in HPC workflows, where users can dynamically interact with the execution of a workflow for purposes such as analysis and debugging.

Next to the presented papers, the workshop featured two invited talks: the first by prof. Paul Watson from Newcastle University, UK on *Realizing the Potential of the Cloud for Workflow: Scalability, Security and Reproducibility*, and the second talk by Jelena Pjesivac-Grbovic from Google on *The Google Cloud Platform* and giving an overview of the various distributed data processing frameworks offered and developed by Google.

Details of the papers, keynotes and tutorials are available on the workshop web-site[1], and the proceedings are published on the ACM DL [1]. The rest of the report provides a summary of the contributions, and is structured along the distinction in scope and purpose introduced above.

2. PAPER PRESENTATIONS

DynamicCloudSim: Simulating Heterogeneity in Computational Clouds

In this paper from the Humboldt-University in Berlin, Marc Nicolas Bux, on behalf of Ulf Leser, presented the *DynamicCloudSim* system. It extends the popular framework CloudSim [3] for sim-

[1] http://sites.google.com/site/sweetworkshop2013

ulating resource provisioning and scheduling algorithms on cloud computing infrastructure with the aspect of instabilities that are common to shared cloud infrastructures. There are several factors of instability that are taken into account. The first is inhomogeneity in the performance of computational resources observed and measured for example by Dejun et al. [6], Jackson et al. [8] and Schad et al. [12]. A second factor is uncertainty in, and dynamic changes to, the performance of VMs due to sharing of common resources with other VMs and users, as for instance reported by Dejun et al. [6]. A third and final factor are straggler VMs [16] and failures during task execution for which programmers need to be prepared for, especially in massively parallel applications [13]. Accounting for those types of instabilities makes simulations of cloud applications more reliable which is important for cost planning. It is also the first step for evaluating novel approaches towards resource allocation and task scheduling on distributed architectures. DynamicCloudSim has been tested on scientific workflow scenarios, yet it is still to be verified against traces of workflow execution on actual cloud infrastructure.

A Continuous Workflow Scheduling Framework

Panayiotis Neophytou from the Department of Computer Science, University of Pittsburgh, Pittsburgh, PA, USA, on behalf of Panos Chrysanthis and Alexandros Labrinidis argued that both scientific and business WfMS can be extended to support data stream semantics to enable monitoring applications. For this goal the authors designed and implemented an integrated workflow scheduling framework STAFiLOS, a STreAm FLOw Scheduler, a Continuous Workflows framework within their CONFLuEnCE [10] engine built on top of the Kepler system [9]. STAFiLOS supports the implementation of different scheduling policies. It was evaluated based on the Linear Road Benchmark [2] — the standard benchmark for stream processing system — and compared against Kepler's own Thread-Based director.

OSIRIS-SR – A Scalable yet Reliable Distributed Workflow Execution Engine

Nenad Stojnic from the Department of Mathematics and Computer Science, University of Basel, Switzerland, presented this paper on behalf of the second author Heiko Schuldt. It introduces OSIRIS-SR (Open Service Infrastructure for Reliable and Integrated process Support – Safety Ring) a true peer-to-peer workflow execution engine, which is an extension of the OSIRIS system [14, 15]. In contrast to other workflow engines, here the workflow orchestration itself is distributed across a set of cooperating nodes. This is done by means of *mini workflow engines* on each node, which together form the *OSIRIS-SR layer*. This results in higher scalability and reliability. To protect against network or node failures OSIRIS-SR uses a scalable self-organizing and self-healing *node monitor* overlay, called the Safety Ring. Its members supervise the non-member nodes currently in charge of service invocation and also provide a scalable and reliable metadata storage. The presented evaluation results show that the Safety Ring-based failure handling and transactional migration at instance level comes with a only minor and affordable impact on the overall performance.

User-Steering on HPC Workflows: State of the Art and Future Directions

Daniel de Oliveira from Fluminense Federal University, Niterói, Brazil summarized the state-of-the-art and the main challenges in supporting user-steering in HPC workflows. The other authors are Marta Mattoso, Jonas Dias, Kary Ocaña, Flavio Costa, Felipe Horta, Vítor Sousa and Igor Araújo from COPPE, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil and Eduardo Ogasawara from Federal Center for Technological Education, Rio de Janeiro, Brazil. User-steering of workflows refers here to the run-time interference of users with the execution of a workflow. This can mean for example stopping the workflow, analysing intermediate results, changing parameters or even the structure of the workflow, and finally letting the execution continue. User-steering is a big step towards more dynamic workflows and fully supporting the exploratory nature of Science and the dynamic process involved in scientific analysis [7]. Based on the motivation from domains of Bioinformatics and the Gas & Oil domain and based on the previous experience such as [11] three main issues were formulated and discussed: (i) monitoring of execution, (ii) data analysis at runtime, and (iii) dynamic interference in the execution. This division guided the discussion of the state-of-the-art in workflow steering. The conclusion was that in (i) the desired features are already tackled by existing support for querying provenance at runtime [11, 5] and scientific event notification [4]. For (ii) several open challenges were discussed as data staging, big data *in situ* analysis, decision-support tools, dynamic workflow engines, parameter slice exploration and experiment optimization.

3. KEYNOTES

Realizing the Potential of the Cloud for Workflow: Scalability, Security and Reproducibility

The workshop started with a first invited talk by Prof. Paul Watson from Newcastle University, UK who's focus were cloud-based workflow systems. The talk discussed the opportunities offered by cloud computing in overcoming some of the limitations of service-based approaches to workflows enactment. Specifically, three areas were discussed where synergies can be found between workflow technology and cloud computing, namely scalability, security, and reproducibility. Exploiting such synergies, however, requires a radical redesign of the workflow management system, rather than simply a porting of existing implementations to the cloud. An example of such design is e-Science Central, an open-source workflow platform developed by the Information Management group at Newcastle University. This WFMS runs natively on multiple public cloud infrastructures, and is aimed at supporting workflows which are deployed over hundred of cloud nodes and with a running time that is measured in weeks, in areas such as chemical engineering (QSAR) and activity recognition for medical applications.

The Google Cloud Platform

The workshop finished with a second invited talk by Dr Jelena Pjesivac-Grbovic from Google, Inc. on The Google Cloud Platform. Jelena Pjesivac-Grbovic is a senior software engineer in Systems Infrastructure at Google, focusing on distributed data processing frameworks. The Google Cloud Platform is a collection of services offered by Google that allows external users to build their applications and run their computations on top of the Google infrastructure. It consists of the following parts: *App Engine* which allows developers to create apps that are easy to manage and scale, *Cloud Datastore* which offers a schema-less, non-relational datastore with built-in query support, *CloudSQL* which lets developers run MySQL databases in Google Cloud, *Compute Engine* which can run large-scale computing workloads on Linux virtual machines, *Cloud Storage* for storing, accessing and managing data, *BigQuery* for interactive analysis of datasets with billions of rows, *Prediction API* for applying machine learning and finally *Translation API* for automatic translation into other languages.

In this talk Jelena focused on the parts of Google Cloud Platform for executing large-scale data-intensive workflows, which are Google App Engine, Google Compute Engine, Google Cloud Storage and Google Big Query. Two use-cases were presented where these services where used in concert to do large scale data collection and processing. The first was to collect and present a queryable visual interface to show the positions of all the ships in the world. The data would be collected by loggers into Cloud Storage, and the interface was built using Cloud Storage and BigQuery. The second use case was a data sensing lab for collecting and analyzing data during the 2012 Google I/O event by "mote" robots and sensors that monitored participants and environmental parameters such as temperature, pressure, humidity, quality of air, light and RF noise. This resulted in more than 10 GB of data per 20 seconds, which was collected using App Engine and Cloud Datastore, and subsequently analyzed using Cloud Storage and BigQuery with the possible additional use of R and Hadoop. Both presented use cases aimed at showing the practicality and scalability of data-intensive workflows built upon the different presented services of the Google Cloud Platform.

4. CONCLUSION

The presentations and tutorials at SWEET 2013 provided an overview of current developments and emerging issues in the area of dynamic workflow execution by which we mean here the type of workflow execution where during the execution there are changes in the input of the workflow, the specification of the workflow or the distributed backend. These proceedings show that although much has been achieved in this area to make large-scale data-intensive computing more robust and practical, there is still much left for further research. Specifically the following issues and topics were raised and discussed during the workshop:

User-friendliness and Workflow Design Assistance: One of the goals of data-intensive workflow systems is to make big-data computing platforms more usable for non-programmers. However, their user-interfaces are up to now still fairly technical and not giving much assistance with designing effective workflows for certain tasks. The interface could for example recommend certain components or patterns, based on a task description, or it could detect anti-patterns that signal an incorrect or inefficient workflow.

Heterogeneous data-processing workflows: In practice workflows often have to process different types of data from different sources to produce the final result. The data sources may differ in data complexity and retrieval speed, but also in whether

the data is retrieved in big chunks or as a stream of small chunks. At the same time the components in the workflow may also be very different, some may be simple arithmetical operations while others are complex database queries. All this makes it harder to efficiently schedule and execute the workflow, and requires additional research.

Realistic performance models for computational clouds: The efficient scheduling and optimisation of data-intensive workflows, both dynamically and statically, depends to a large extent on having realistic and reliable models for estimating the cost of a schedule or evaluation plan. Research is needed into which types of computational clouds that currently are offered have which types of performance characteristics, and how reliably their behavior can be predicted for the purpose of optimization.

As can be seen from this list, there is no lack of research challenges for the future editions of the SWEET workshop, and interesting papers investigating them are therefore to be expected.

Acknowledgements: We would like to thank the PC members, keynote speakers, authors, local workshop organizers and attendees for making SWEET 2013 a successful workshop. We also express our great appreciation for the support from Google Inc.

5. REFERENCES

[1] *SWEET '13: Proceedings of the 2nd ACM SIGMOD Workshop on Scalable Workflow Execution Engines and Technologies*, New York, NY, USA, 2013. ACM.

[2] Arvind Arasu, Mitch Cherniack, Eduardo Galvez, David Maier, Anurag S. Maskey, Esther Ryvkina, Michael Stonebraker, and Richard Tibbetts. Linear road: a stream data management benchmark. In *Proceedings of the Thirtieth international conference on Very large data bases - Volume 30*, VLDB '04, pages 480–491. VLDB Endowment, 2004.

[3] Rodrigo N. Calheiros, Rajiv Ranjan, César A. F. De Rose, and Rajkumar Buyya. Cloudsim: A novel framework for modeling and simulation of cloud computing infrastructures and services. *CoRR*, abs/0903.2525, 2009.

[4] Flavio Costa, Vítor Silva, Daniel de Oliveira, Kary Ocaña, Eduardo Ogasawara, Jonas Dias, and Marta Mattoso. Capturing and querying workflow runtime provenance with prov: a practical approach. In *Proceedings of the Joint EDBT/ICDT 2013 Workshops*, EDBT '13, pages 282–289, New York, NY, USA, 2013. ACM.

[5] D. de Oliveira, E. Ogasawara, F. Baião, and M. Mattoso. Scicumulus: A lightweight cloud middleware to explore many task computing paradigm in scientific workflows. In *Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on*, pages 378–385, 2010.

[6] Jiang Dejun, Guillaume Pierre, and Chi-Hung Chi. Ec2 performance analysis for resource provisioning of service-oriented applications. In *Proceedings of the 2009 international conference on Service-oriented computing*, ICSOC/ServiceWave'09, pages 197–207, Berlin, Heidelberg, 2009. Springer-Verlag.

[7] Y. Gil, E. Deelman, M. Ellisman, T. Fahringer, G. Fox, D. Gannon, C. Goble, M. Livny, L. Moreau, and J. Myers. Examining the challenges of scientific workflows. *Computer*, 40(12):24–32, 2007.

[8] Keith R. Jackson, Lavanya Ramakrishnan, Krishna Muriki, Shane Canon, Shreyas Cholia, John Shalf, Harvey J. Wasserman, and Nicholas J. Wright. Performance analysis of high performance computing applications on the amazon web services cloud. In *Cloud Computing, Second International Conference, CloudCom 2010, November 30 - December 3, 2010, Indianapolis, Indiana, USA, Proceedings*, pages 159–168, 2010.

[9] Bertram Ludäscher, Ilkay Altintas, Chad Berkley, Dan Higgins, Efrat Jaeger, Matthew Jones, Edward A. Lee, Jing Tao, and Yang Zhao. Scientific workflow management and the kepler system: Research articles. *Concurr. Comput. : Pract. Exper.*, 18(10):1039–1065, August 2006.

[10] Panayiotis Neophytou, Panos K. Chrysanthis, and Alexandros Labrinidis. Confluence: Implementation and application design. In Dimitrios Georgakopoulos and James B. D. Joshi, editors, *7th International Conference on Collaborative Computing: Networking, Applications and Worksharing, CollaborateCom 2011, Orlando, FL, USA, 15-18 October, 2011*, pages 181–190, 2011.

[11] Eduardo S. Ogasawara, Daniel de Oliveira, Patrick Valduriez, Jonas Dias, Fabio Porto, and Marta Mattoso. An algebraic approach for data-centric scientific workflows. *PVLDB*, 4(12):1328–1339, 2011.

[12] Jörg Schad, Jens Dittrich, and Jorge-Arnulfo Quiané-Ruiz. Runtime measurements in the cloud: observing, analyzing, and reducing variance. *Proc. VLDB Endow.*,

3(1-2):460–471, September 2010.

[13] Bianca Schroeder and Garth A. Gibson. A large-scale study of failures in high-performance computing systems. In *Proceedings of the International Conference on Dependable Systems and Networks*, DSN '06, pages 249–258, Washington, DC, USA, 2006. IEEE Computer Society.

[14] Christoph Schuler, Heiko Schuldt, Can Türker, Roger Weber, and Hans-Jörg Schek. Peer-to-peer execution of (transactional) processes. *Int. J. Cooperative Inf. Syst.*, 14(4):377–406, 2005.

[15] Christoph Schuler, Can Turker, Hans-Jorg Schek, Roger Weber, and Heiko Schuldt. Scalable peer-to-peer process management. *International Journal of Business Process Integration and Management*, 1(2):129–142, 2006.

[16] Matei Zaharia, Andy Konwinski, Anthony D. Joseph, Randy Katz, and Ion Stoica. Improving mapreduce performance in heterogeneous environments. In *Proceedings of the 8th USENIX conference on Operating systems design and implementation*, OSDI'08, pages 29–42, Berkeley, CA, USA, 2008. USENIX Association.