IJRRECS/August 2013/Volume-1/Issue-4/332-337 ISSN 2321-5461

INTERNATIONAL JOURNAL OF REVIEWS ON RECENT ELECTRONICS AND COMPUTER SCIENCE

ADVANCEMENT TOWARDS THE CHALLENGES OF SCIENTIFIC WORKFLOW

Govardhan Reddy[1], G.Shiva Krishna[2]

[1]M.Tech Student, Dept of CSE, RRS College of Engineering and Technology, Patancheru, Medak, A.P, India

[2]Associate Professor, Dept of CSE, RRS College of Engineering and Technology, Patancheru, Medak, A.P, India

ABSTRACT:

With the rising complexity, the development of modern science has created entire quantity of data. Processing and supervision such large-scale scientific data sets are usually beyond the realms of individual scientists to solve; instead, it has to rely on multiple domain scientists with diverse expertise. Such scientific data analysis and processing is usually structured and automated by a dataflow-oriented process called scientific workflow. However, such workflows have to be intended and improved among numerous scientists over a long time period. In contrast to business processes that are control-flow oriented and arrange a collection of well-defined business tasks to attain a business goal, technical workflows are often dataflow-oriented and reorganize a collection of technical tasks to allow and accelerate technical detection. Accessible tools are single user-oriented and do not sustain workflow expansion in a collaborative fashion. Accessible technical workflow tools do not mainly support collaborative composition. Attribution has been widely considered critical to the reproducibility of scientific workflows. Compared to existing considerable amount of work focusing on attribution for run-time workflow implementation, our work focuses on association attribution that tracks human connections and efforts in the process of technical workflow composition.

Keywords: Data Flow Oriented Process, Scientific Workflow, Collaborative Fashion, And Single User Orientation.

 ISSN 2321-5461

1. INTRODUCTION:

The improvement of present knowledge has formed sheer volume of data with mounting complication. Scientific data examination and handing out is usually structured and programmed by a dataflow-oriented process unrestricted scientific workflow [1]. Technical workflows are habitually dataflow-oriented and streamline a collection of scientific tasks to enable and accelerate scientific discovery in contrast to business processes that are control-flow oriented and orchestrate a collection of well-defined business tasks to achieve a business goal. To integrate and structure local and remote heterogeneous computational and data resources to perform in silico experiments scientific workflows are used [3]. The increasingly important responsibility of scientific workflows in current science was highlight in which the speediness with which any given discipline move ahead is likely to depend on how well the community get hold of the essential expertise in database, workflow organization, revelation, and cloud computing knowledge. Scientific workflow and scientific collaboration are two means method to hold up scientific data analysis and organization.

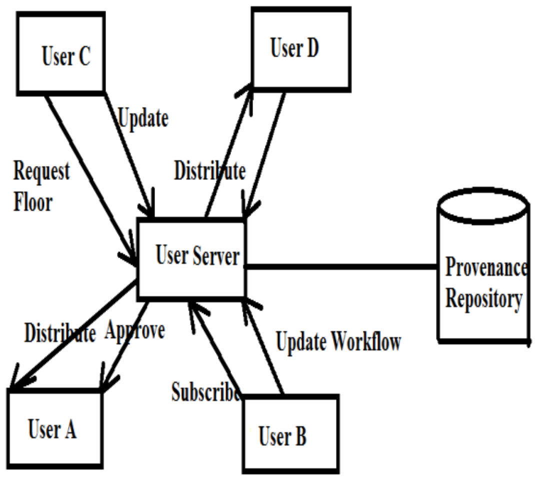

Fig 1: An overview of Collaborative Composition Replica.

The junction of the two trends obviously leads to a perception that we coined as collaborative scientific workflow meaning that multiple scientists are complicated and pool resources in the complete lifecycle of a work- flow its design, modification, implementation, supervise, and administration [2]. Collaborative workflow shown in fig1 proposes is typically critical for the achievement of a comprehensive workflow work of art. Very easy Large Synoptic Survey Telescope experiments represent a three-step workflow data recovery, pre-processing, and data representation. Scheming each step necessitate different expertise. In the intervening time, the consecutive relationship among the steps involve that the accurateness of the complete process is resolute by the proposal of each step and the error proliferation between them [4] [6]. While concerning scientists with different potential focus on discovery a local optimal design of a meticulous step, association between them will find a comprehensive optimal design for the complete workflow. Active scientific workflow tools do not for the most part support collaborative composition. Conveniences that maintain scientists in collaboratively designing workflows are limited. Individual work artefacts are physically sent to associate or uploaded to some mutual social space to facilitate collaborative design and conversation [5]. Such a detached collaboration style apparently does not sustain real-time discussion and collaboration, which are more often than not critical to scientific examination. In addition, provenance data is not preserve adequately, which is significant for workflow reproducibility [7]. Hence a design environment, capable with system-level support for collaborative scientific workflow composition is projected.

2. MODELLING OF BUSINESS WORKFLOWS:

The term collaborative work- flows is compatible with the term coordinated work- flows for business workflows which put emphasis on harmonization between workflows. The term is used to give emphasis to the contribution of multiple scientists with a workflow. A number of works has been carrying out to help model and arrange business workflows. Decentralized services- oriented middleware architecture was to compose workflows and identify their dependencies [8]. A two-layer framework was proposed where Web services are modelled as self-coordinating unit and a workflow communicate such entities into a network of objects. A verbal communication to identify workflow composition, encompass both reactive tasks and proactive tasks in a Grid environment was proposed [10]. Workflow harmonization as molecules and reactions to facilitate autonomous evolution in a changing Grid environment was proposed. In difference to business process modelling where encompassing mechanism is openly connected all the way through control links, tasks in methodical workflow modelling are connected all the way through data links, which have to be measured when locks are assigning to tasks during concerted workflow design [9] [11]. A concept of harmonization area considering pre- or post-conditions for each comprising task was proposed and use event-condition-action rules to identify some runtime behaviour at plan time.

3. PRESERVATION OF COLLABORATION GROUPING INFORMATION BY THE CENTRAL SERVER:

Scientific collaborations frequently last for an extensive period of time. In addition, provisional discussion groups and sessions may be produced in the lifecycle of a long-term scientific cooperation process and hence a hierarchical structure for the central server is proposed [12]. It may host numerous collaboration groups, which may or may possibly not have nesting associations between them. The central server preserves all collaboration groups in- formation and takes action as the subject for all inventory groups. All observers are controlled into corresponding association groups. The central server also stores and administers all attribution data, so that it becomes ordnance of workflow commodities and facilitates scalability [14] [16]. Within a collaboration group, an uncomplicated way is to permit everyone to do anything on a workflow at several times, and hand out the results to everyone in the similar group. Normally only one person is authorized to speak at a certain moment in a group consequently, we grant right of entry control policies so that only one person at a time can adjust the shared workflow products and share out the changes in the group [13]. The floor control technique is implemented from an expansively tested and well proved human communication protocol, where a single floor is preserved in a shared meeting background. Every member requests and contends for the floor, and only the person who gets hold of the floor can converse in the meeting. The people who schedule an association group at the central server become the mediator of the group, and will involuntarily have the control over the floor [15]. Multiple floor-based access control facility, by means of finer-grained locking mechanism is used for superior concurrency.

4. CONCLUSION:

The ongoing work on setting up collaboration protocols to maintain collaborative technical workflow composition is presented in our paper. Our service-oriented infrastructure includes a collaboration ontology connected with a set of collaboration prototypes, primitives, and builds, as well as synchronized control mechanisms to maintain synchronized collaborative workflow composition. We plan to continue our investigation in the following directions. First, we will propose and accomplish an evaluation study and use the response to improve the system. Second, we plan to improve collaboration attribution organization performance based on the collaboration ontology. Third, we plan to perform more experiments to learn the effects of tuning a variety of parameters on synchronized productivity.

REFERENCES:

[1] G.M. Olson, A. Zimmerman, and N. Bos, eds., Scientific Collaboration on the Internet , 2008, MIT Press, Cambridge, MA, USA.

[2] LSST, "Large Synoptic Survey Telescope", 2009, Accessed on, Available from: http://www.lsst.org/lsst/science .

[3] B. Ludäscher, "Scientific Workflows: Cyberinfrastructure for e- Science", in Proceedings of Pacific Neighborhood Consortium (PNC) , 2007, Berkeley, CA, USA, Oct. 19, pp.

[4] Y. Gil, E. Deelman, J. Blythe, C. Kesselman, and H. Tangmunarunkit, "Artificial Intelligence and Grids: Workflow Planning and Beyond", IEEE Intelligent Systems , Jan.-Feb., 2004, 19(1): pp. 26–33.

[5] E. Deelman and Y. Gil, "NSF Workshop on the Challenges of Scientific Workflows", (ed.), May 1-2, 2006.

[6] S. Wuchty, B. Jones, and B. Uzzi, "The Increasing Dominance of Teams in Production of Knowledge", Science , 2007, 316: pp. 1036- 1039.

[7] N.R. Council, "Facilitating Interdisciplinary Research". 2004, National Academies Press, Washington DC, USA.

[8] G. Bell, T. Hey, and A. Szalay, "Beyond the Data Deluge", Science , 2009, 323(5919): pp. 1297-1298.

[9] S. Lu and J. Zhang, "Collaborative Scientific Workflows Supporting Collaborative Science", International Journal of Business Process Integration and Management (IJBPIM) , 2011, 5(2): pp. 185-199.

[10] B. Ludäscher, I. Altintas, C. Berkley, D. Higgins, E. Jaeger, M. Jones, E.A. Lee, J. Tao, and Y. Zhao, "Scientific Workflow

Management and the Kepler System", Concurrency and Computation: Practice & Experience , 2006, 18(10): pp. 1039-1065.

[11] T. Oinn, M. Greenwood, M. Addis, M.N. Alpdemir, J. Ferris, K. Glover, C. Goble, A. Goderis, D. Hull, D. Marvin, P. Li, P. Lord, M.R. Pocock, M. Senger, R. Stevens, A. Wipat, and C. Wroe, "Taverna: Lessons in Creating a Workflow Environment for the Life Sciences", Concurrency and Computation: Practice & Experience , 2006, 18(10): pp. 1067–1100.

[12] D. Churches, G. Gombas, A. Harrison, J. Maassen, C. Robinson, M. Shields, I. Taylor, and I. Wang, "Programming Scientific and Distributed Workflow with Triana Services", Concurrency and Computation: Practice & Experience , 2006, 18(10): pp. 1021–1037.

[13] J. Freire, C.T. Silva, S.P. Callahan, E. Santos, and C.E. Scheidegger, "Managing Rapidly-Evolving Scientific Workflows ", Lecture Notes in Computer Science , May, 2006, 4145/2006: pp. 10–18.

[14] Y. Zhao, M. Hategan, B. Clifford, I. Foster, G. Laszewski, V. Nefedova, I. Raicu, T. Stef-Praun, and M. Wilde, "Swift: Fast, Reliable, Loosely Coupled Parallel Computation", in Proceedings of IEEE International Workshop on Scientific Workflows, 2007, Salt Lake City, UT, USA, Jul. 9-13, pp. 199–206.

[15] A. Chebotko, C. Lin, X. Fei, Z. Lai, S. Lu, J. Hua, and F. Fotouhi, "VIEW: A Visual Scientific Workflow Management System", in Proceedings of the 1st IEEE International Workshop on Scientific Workflows , 2007, Salt Lake City, UT, USA, Jul., pp. 207–208.

[16] C. Lin, S. Lu, Z. Lai, A. Chebotko, X. Fei, J. Hua, and F. Fotouhi, "Service-Oriented Architecture for VIEW: A Visual Scientific Workflow Management System", in Proceedings of the IEEE 2008 International Conference on Services Computing (SCC) , 2008, Honolulu, HI, USA, Jul. 9-11, pp. 335-342.