Web Scale Trustworthy Collaborative Service Systems The Coast Project

François Charoy and the members of the Coast Team | September 2014

FIELD: Networks, Systems and Services, Distributed Computing

CONTENTS

1 TEAM COMPOSITION

PERMANENT PEOPLE INVOLVED IN THE PROJECT

- François Charoy, Professor, HDR, Université de Lorraine, Team Leader.
- Khalid Benali, Associate Professor, HDR, Université de Lorraine.
- Gérome Canals, Associate Professor, HDR, Université de Lorraine.
- Claude Godart, Professor, HDR, Université de Lorraine.
- Claudia Ignat, Inria Researcher.
- Gérald Oster, Associate Professor, Université de Lorraine.
- Olivier Perrin, Professor, HDR, Université de Lorraine.
- Pascal Urso, Associate Professor, Université de Lorraine.
- Samir Youcef, Associate Professor, Université de Lorraine.

RESEARCH FELLOWS

PHD STUDENTS

- Luc André, PhD Student, Université de Lorraine.
- Ahmed Bouchami, PhD Student, Université de Lorraine.
- Elio Goettelmann, PhD Student, Université de Lorraine, CRP Henri Tudor.
- Jordi Martori, PhD Student, Université de Lorraine, Inria.
- Mehdi Ahmed Nacer, PhD Student, Université de Lorraine.

INVITED PROFESSORS

2 VISION & OBJECTIVES

The advent of the Cloud, of smart mobile devices and of service-based architecture has opened a field of possibilities as wide as the invention of the Web 20 years ago. With Web 2.0, software companies can deliver applications and services using the web as a platform. From text to video editing, from data analytics to process management, they can distribute business applications to users on their web browser or on some mobile appliance [1]. These services are mostly human-centred and deployed on sophisticated infrastructures that can cope with very high loads. The Software as a Service approach (SaaS) highlights their cooperative nature, by enabling the storage of data in cloud infrastructures that can be easily shared among users. Thus, clients consume applications through service API (web services), available on delivery platforms, called stores or markets. This approach of the distribution of software outstrips the traditional software distribution channels, in both scale and opportunity. Here, scale concerns the number of users (communities rather than groups), the size of data produced and managed (billions of documents), the number of services and of organizations (tens of thousands). Opportunity refers to the infinite number of combinations between these services.

Just like the web, evolution occurred very quickly and continues at a fast pace. It challenges research because the creation of applications from the composition of services must incorporate

1. See http://blog.programmableweb.com/2011/09/16/open-api-growth-a-visualization/

new content and context based constraints. From a socio-technical perspective, the behaviour of users is evolving constantly. Mere enhancement of current existing solutions to cope with these challenges is likely insufficient. We propose a dedicated research effort to tackle the problems arising from the evolution of contemporary technologies.

For this purpose, we propose a framework to explore three directions: large scale collaborative data management, data centred service composition and above all, a foundation for the construction of trustworthy collaborative systems.

Large scale collaborative data management concerns mostly the problem of allowing people to collaborate on shared data, synchronously or not, on a central server or on a peer to peer network. Although this research has a long history, new challenges arise regarding needs that are occurring with the acculturation of users to collaboration like the number of participants to a collaboration (a crowd), sharing among different organisations and the nature of document that are shared and produced. The problem here is to design new algorithms and to evaluate them under different usage conditions and constraints and for different kinds of data. This research is in the continuation of the ECOO and SCORE projects, with new perspectives on scale and the methodological approach. It is described in section 4.1.

Data centred service composition deals with the challenge of creating applications by composing services from different providers. Service composition has been studied for some time now but the technical evolution and the growing availability of public API oblige us to reconsider its problematic [6]. Our goal here is, taking into account this evolution, like the advent of the Cloud, the availability at a large scale of public API based on the REST architectural style, to design models, methods and tools to help developers to compose these services in a safe and effective way. Again this research is a continuation of the previous results of ECOO and SCORE on services taking into account the evolution of our landscape. It is described in section 4.2

Based on the work that we have done and that we will conduct in the two first directions, our main research direction aims at providing support to build trustworthy collaborative applications based on the knowledge that we can gather from the underlying algorithms, from the composition of services and from services quality that can be deduced and monitored. The complexity of the context in which applications are executed does not allow to provide proven guarantees. Our goal is to base our work on a contractual and monitored approach to give to users confidence in the service they use. It is very surprising to see how people rely today on services with very little knowledge about the amount of confidence they put in these services. As soon as these services are based on composition of other unknown services, it becomes very difficult to understand the consequences of the failure of a component of the composition for instance. This direction is our most ambitious and long term endeavour for this project, described in section 4.3.

Figure 1 shows the path from the ECOO and SCORE research topics to the Coast project. The path portrays a ruptured continuum, to underscore both the endurance of the common questions along with the challenge of accommodating a new scale. We regard collaborative systems as a combination of supportive services, encompassing safe data management and data sharing. Trustworthy data centred services are essential support for collaboration at the scale of communities and organisations. We see there that we aim at combining our results and expertise to achieve a new leap forward toward the understanding and the mastering of methods and techniques that allow the engineering and the use of large scale collaborative systems.

---

[6] Fabio Casati. Promises and failures of research in dynamic service composition. In Janis Bubenko, John Krogstie, Oscar Pastor, Barbara Pernici, Colette Rolland, and Arne Sølvberg, editors, *Seminal Contributions to Information*

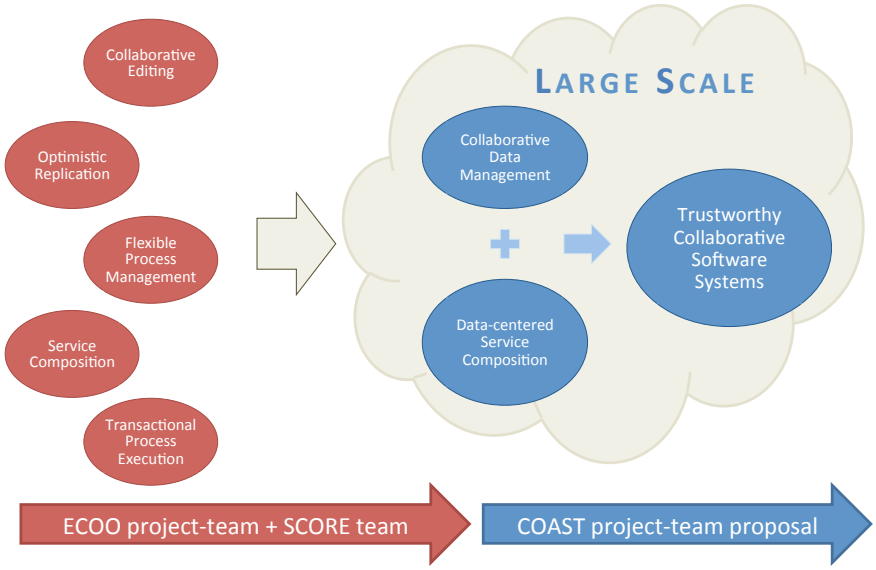

Figure 1: Coast Research Directions

Before describing our project in more detail, we review the achievements of the ECOO and SCORE teams, that serve as our foundation.

3 PREVIOUS RESULTS AND ACHIEVEMENTS

The proposed project extends the work of the ECOO project (1998-2009) and SCORE team (2009-2014) conducted over more than ten years in the area of collaboration between people, systems and organisations. The ECOO Project (Environment and Cooperation) was lead by Claude Godart until october 2009. It was followed by the SCORE (Service and Cooperation) team lead by Pascal Molli until september 2010 and François Charoy since then.

We considered that collaboration occurs when participants can share data and coordinate operations on these data. Thus, the dimensions of sharing and coordination structured our work.

Sharing allows people to cooperate by exchanging data. People coordinate their work implicitly using knowledge of other participants' actions. Our main use case concerns the implicit coordination in collaborative editing over sets of shared documents. Coordination relies on shared mental models of the editing task for those documents. Application examples include Wikipedia, real-time editing like GoogleDocs or version control systems.

Users need awareness mechanisms to reduce the number of conflicts from concurrent document editing. To address these concerns, our team has advanced optimistic replication algorithms using operational transformation, designed a new class of algorithms based on commutativity (also called CRDT) and provided novel awareness support.

Systems Engineering, pages 235–239. Springer Berlin Heidelberg, 2013.

Cooperation also requires the explicit modelling of coordination as an orchestration of activities and service calls. Our approach to modelling also accommodates non functional properties of orchestration such as transactional behaviour, and issues of security and temporal relationships. The distributed nature of the services required research in the distribution of their execution, drawing on traditional orchestration as well as ad-hoc methods for special contexts such as crisis management.

Our results have been validated theoretically and experimentally, with scientific papers publications, a patent and by the development and the transfer of various software tools (Bonita, LibreSource, QualiPSo Factory, XWiki 3.0).

We are at present positioning our work with a a longer term perspective. We will extend our understanding of computer supported collaboration to accommodate the new scale offered by the evolution of low cost functions (mobile, ubiquitous, always connected, on multiple devices) and IT infrastructure (service architecture, the public/private/hybrid clouds, the various service delivery models, peer-to-peer networks).

The team, in its succcessive configurations, has contributed to the design of the main families of optimistic replication algorithms for maintaining consistency: operational transformation (OT) and commutative replicated data types (CRDT). The Score team pioneered research on CRDT algorithms that ensure consistency of highly dynamic content on peer-to-peer networks. These CRDT algorithms rely on natively commutative operations defined on abstract data types such as lists or ordered trees. We first developed WOOT and WOOTO algorithms [53, 30]. We also developed the LOGOOT [56, 55, 52] and the CRDT approach that overcomes the WOOT constraint of retaining information about deleted elements. Finally, we proposed a novel undo approach using a compensation mechanism [54, 56].

We have also succeeded in adapting these algorithms to handle complex data in peer-to-peer environments. A principled solution of the reconciliation of file systems and hierarchical documents such as XML trees appears in [40, 37, 38, 34, 32]. Other types of complex data that we investigated include peer-to-peer wikis [16, 15, 47, 46, 45] and distributed semantic wikis [43]. We focused on an approach for the distributed construction of semantic wiki pages using the publish-subscribe model where the publication, propagation and integration of modifications are controlled by users [42, 29, 41]

We investigated an awareness mechanism for improving collaborative work in the context of software engineering [33, 31], collaborative editing of textual documents [36], web pages [44] and peer-to-peer wikis [4, 5]. Moreover, in [18, 19] we tackled the challenge of divergence awareness in collaborative knowledge management.

We have examined unique ways to support the development of safe and secure programs as a composition of services. Regarding the safety of the service composition, we leveraged a declarative approach and model checking to describe, enforce and integrate non functional properties in service composition. These ranged from time properties attached to service composition [28] to transactional properties of service composition [14] and to security properties [61, 8]. Following this track, we have also made significant contributions in the domain of service engineering, where classical software engineering problems resurface with a stronger focus on reuse, security and compliance [27, 48, 59]. Recently, we have pursued the distributed orchestration of services [57, 20, 58, 39]. One of our challenges is to apprehend the relationships between service composition, its security constraints, the distribution of its execution and the underlying component architecture [60, 17].

At the frontier between collaborative systems and process modelling, mainstream research

acknowledges the need to introduce flexibility at different levels in the process model design and execution. SAP Research funded our work on the use of delegation as a means to introduce flexibility at the organisational level. This led us to establish the relationship between delegation as part of a day to day organisational policy for process management and as part of a security policy [24]. We also explored the application of traditional process management to the coordination of crisis management activities. This led us to a novel proposition for inter-organisational coordination, that departs from traditional models, and integrated our work in distributed collaborative systems [21]

Now we must incorporate the global evolution of computing in our research, to understand and anticipate its impact on our vision and its evolution.

TEAM REFERENCES

[1] Mehdi Ahmed-Nacer, Pascal Urso, and François Charoy. Improving Textual Merge Result. In *9th IEEE International. Conference on Collaborative Computing: Networking, Applications and Worksharing*, Austin, Texas, États-Unis, 2013.

[2] Mehdi Ahmed-Nacer, Pascal Urso, and François Charoy. Evaluating Software Merge Quality. In *18th International Conference on Evaluation and Assessment in Software Engineering*, London, Royaume-Uni, March 2014. ACM.

[3] Mehdi Ahmed-Nacer, Pascal Urso, Balegas Valter, and Nuno Preguiça. Concurrency Control and Awareness Support for Multi-synchronous Collaborative Editing. In *9th IEEE International. Conference on Collaborative Computing: Networking, Applications and Worksharing*, Austin, Texas, États-Unis, 2013.

[4] Sawsan Alshattnawi. *Concurrence et Conscience de Groupe dans l'Édition Collaborative sur Réseaux Pair-à-Pair.* PhD thesis, Université Henri Poincaré - Nancy I, 11 2008.

[5] Sawsan Alshattnawi, Gérôme Canals, and Pascal Molli. Concurrency Awareness in a P2P Wiki System. In Waleed w. Smari and William McQuay, editors, *2008 International Symposium on Collaborative Technologies ans Systems - CTS 2008*, Irvine, CA, USA, 2008. IEEE Press.

[6] Luc André, Stéphane Martin, Gérald Oster, and Claudia-Lavinia Ignat. Supporting Adaptable Granularity of Changes for Massive-scale Collaborative Editing. In *CollaborateCom - 9th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing - 2013*, Austin, États-Unis, 2013.

[7] Aymen Baouab, Walid Fdhila, Olivier Perrin, and Claude Godart. Towards Decentralized Monitoring of Supply Chains. In *The 2012 IEEE Nineteenth International Conference on Web Services (ICWS 2012)*, Honolulu, États-Unis, 2012.

[8] Aymen Baouab, Olivier Perrin, Nicolas Biri, and Claude Godart. Security Meta-Services Orchestration Architecture. In *IEEE Asia-Pacific Services Computing Conference - APSCC 2009*, Biopolis, Singapour, 2009. IEEE.

[9] Aymen Baouab, Olivier Perrin, and Claude Godart. An Optimized Derivation of Event Queries to Monitor Choreography Violations. In C Liu, H Ludwig, F. Toumani, and Q. Yu, editors, *10th International Conference on Service Oriented Computing (ICSOC 2012)*, volume 7636, pages 222–236, Shanghai, Chine, 2012.

[10] Karim Benouaret, Raman Valliyur-Ramalingam, and François Charoy. Answering Complex Location-Based Queries with Crowdsourcing. In *9th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing*, Austin, États-Unis, 2013.

[11] Karim Benouaret, Raman Valliyur-Ramalingam, and François Charoy. CrowdSC: Building Smart Cities with Large-Scale Citizen Participation. *IEEE Internet Computing*, 17(6):57–64, 2013.

[12] Kahina Bessai, Samir Youcef, Claude Godart, and Ammar Oulamara. Bi-criteria strategies for business processes scheduling in cloud environments with fairness metrics. In *RCIS - IEEE Seventh International Conference on Research Challenges in Information Science - 2013*, pages 1–10, Paris, France, 2013.

[13] Kahina Bessai, Samir Youcef, Ammar Oulamara, Claude Godart, and Selmin Nurcan. Resources allocation and scheduling approaches for business process applications in Cloud contexts. In *CloudCom 2012 - 4th IEEE International Conference on Cloud Computing Technology and Science*, Taipei, Taïwan, 2012.

[14] Sami Bhiri, Claude Godart, and Olivier Perrin. Patrons transactionnels pour assurer des compositions fiables de services web. *Technique et Science Informatiques*, 28, 2009.

[15] Gérôme Canals, Pascal Molli, Maire Julien, Stéphane Laurière, Esther Pacitti, and Mounir Tlili. XWiki concerto: un wiki sur réseau P2P supportant le nomadisme. In Emmanuel Dubois and Jean-Marc Pierson, editors, *Proceedings of the 4th French-speaking conference on Mobility and ubiquity computing - UbiMob'08*, volume 277 of *ACM International Conference Proceeding Series*, Saint Malo, France, 2008. ACM.

[16] Gérôme Canals, Pascal Molli, Julien Maire, Stéphane Laurière, Esther Pacitti, and Mounir Tlili. XWiki Concerto: A P2P Wiki System Supporting Disconnected Work. In Springer Berlin / Heidelberg, editor, *5th International Conference, Cooperative Design, Visualization, and Engineering - CDVE'08*, volume 5220/2008 of *Lecture Notes in Computer Science*, Mallorca, Spain, 2008.

[17] Karim Dahman, François Charoy, and Claude Godart. Generation of Component Based Architecture from Business Processes: Model Driven Engineering for SOA. In *ECOWS 2010 - The 8th IEEE European Conference on Web Services Welcome*, Ayia Napa, Greece, 2010.

[18] Farhad Daneshgar, Gérôme Canals, and Alicia Diaz. An Awareness Framework for Divergence Knowledge Communities. In Ned F. Kock, editor, *Encyclopedia of E-Collaboration*. Information Science Reference, 2007.

[19] Alicia Diaz and Gérôme Canals. The collaborative knowledge sharing framework. In *International IADIS International Conference e-Society - IADIS 2007*, Lisbon, Portugal, 2007.

[20] Walid Fdhila and Claude Godart. Toward synchronization between decentralized orchestrations of Composite Web Services. In *5th International Conference on Collaborative Computing: Networking, Applications and Worksharing - CollaborateCom 2009*, Washington, DC, USA, 2009. IEEE.

[21] Jörn Franke, François Charoy, and Paul El Khoury. Collaborative Coordination of Activities with Temporal Dependencies. In *18th International Conference on Cooperative Information Systems - COOPIS'2010*, Crete, Greece, 2010.

[22] Jörn Franke, François Charoy, and Cédric Ulmer. Coordination and situational awareness for inter-organizational disaster response. In *10th International IEEE Conference on Technologies for Homeland Security*, Boston, MA, USA, 2010.

[23] Khaled Gaaloul, François Charoy, and Andreas Schaad. Modelling Task Delegation for Human-Centric eGovernment Workflows. In Soon Ae Chun, Rodrigo Sandoval, and Priscilla Regan, editors, *Proceedings of the 10th Annual International Conference on Digital Government Research: Social Networks: Making Connections between Citizens, Data and Government*, volume 390 of *ACM International Conference Proceeding Series*, Puebla, Mexico, 2009. ACM.

[24] Khaled Gaaloul, Ehtesham Zahoor, François Charoy, and Claude Godart. Dynamic Authorisation Policies for Event-based Task Delegation. In *The 22nd International Conference on Advanced Information Systems Engineering - CAiSE'10*, Hammamet, Tunisia, 2010.

[25] Elio Goettelmann, Walid Fdhila, and Claude Godart. Partitioning and Cloud Deployment of Composite Web Services under Security Constraints. In *IC2E (IEEE International conference on Cloud Engineering)*, San Francisco, États-Unis, 2013.

[26] Elio Goettelmann, Nicolas Mayer, and Claude Godart. A general approach for a trusted deployment of a business process in clouds. In *Medes (Fifth International Conference on Management of Emergent Digital EcoSystems)*, pages 92–99, Luxembourg, Luxembourg, 2013.

[27] Nawal Guermouche, Olivier Perrin, and Christophe Ringeissen. A Methodology For Web Services Composition. Interne, INRIA, 2007.

[28] Nawal Guermouche, Olivier Perrin, and Christophe Ringeissen. Timed Specification For Web Services Compatibility Analysis. *Electronic Notes in Theoretical Computer Science*, 200, 2008.

[29] Hala Skaf Molli, Gérôme Canals, and Pascal Molli. DSMW: Distributed Semantic MediaWiki. In Lora Aroyo, Grigoris Antoniou, Eero Hyvönen, Annette ten Teije, Heiner Stuckenschmidt, Liliana Cabral, and Tania Tudorache, editors, *7th Extended Semantic Web Conference (ESCW 2010)*, volume 6089 of *Lecture Notes in Computer Science*, Heraklion, Greece, 2010. Springer.

[30] Claudia Ignat, Gérald Oster, Pascal Molli, Michèle Cart, Jean Ferrié, Anne-Marie Kermarrec, Pierre Sutra, Marc Shapiro, Lamia Benmouffok, Jean-Michel Busca, and Rachid Guerraoui. A Comparison of Optimistic Approaches to Collaborative Editing of Wiki Pages. In *3rd International Conference on Collaborative Computing: Networking, Applications and Worksharing - CollaborateCom'2007 Proceedings of the The 3rd International Conference on Collaborative Computing: Networking, Applications and Worksharin - CollaborateCom*, White Plains, NY, USA, 2007. IEEE Computer Society.

[31] Claudia-Lavinia Ignat. Annotation of Concurrent Changes in Collaborative Software Development. In *International Conference on Intelligent Computer Communication and Processing - ICCP 2008*, Cluj-Napoca, Romania, 2008.

[32] Claudia-Lavinia Ignat and Moira C. Norrie. Multi-level editing of hierarchical documents. *Journal of Computer Supported Cooperative Work*, 17, 2008.

[33] Claudia-Lavinia Ignat and Gérald Oster. Awareness of Concurrent Changes in Distributed Software Development. In Robert Meersman and Zahir Tari, editors, *International Conference on Cooperative Information Systems - CoopIS 2008 On the Move to Meaningful Internet Systems: OTM 2008*, volume 5331 of *Lecture Notes in Computer Science*, Monterrey, Mexico, 2008. Springer Berlin / Heidelberg.

[34] Claudia-Lavinia Ignat and Gérald Oster. Peer-to-peer Collaboration over XML Documents. In Yuhua Luo, editor, *International Conference on Cooperative Design, Visualization and Engineering - CDVE 2008 Cooperative Design, Visualization, and Engineering*, volume 5220 of *Lecture Notes in Computer Science*, Mallorca, Spain, 2008. Springer Berlin / Heidelberg.

[35] Claudia-Lavinia Ignat, Gérald Oster, Meagan Newman, Valerie Shalin, and François Charoy. Measurement of Remote Response Delay in Multi-Synchronous Collaborative Editing. Rapport de recherche, 2013.

[36] Claudia-Lavinia Ignat, Stavroula Papadopoulou, Gérald Oster, and Moira C. Norrie. Providing awareness in multi-synchronous collaboration without compromising privacy. In ACM, editor, *ACM 2008 conference on Computer supported cooperative work Proceedings of the ACM 2008 conference on Computer supported cooperative work*, San Diego, CA, USA, 2008. ACM.

[37] Stéphane Martin, Mehdi Ahmed-Nacer, and Pascal Urso. Controlled conflict resolution for replicated document. In *Controlled conflict resolution for replicated document*, Pittsburgh, Pennsylvania, États-Unis, 2012.

[38] Stéphane Martin, Pascal Urso, and Stéphane Weiss. Scalable XML Collaborative Editing with Undo short paper. In *CoopIS 2010 18th International Conference on Cooperative Information System - CoopIS'2010*, Crete, Greece, 2010.

[39] Eric Monfroy, Olivier Perrin, and Christophe Ringeissen. Dynamic Web Services Provisioning with Constraints. In Robert Meersman and Zahir Tari, editors, *International Conference on Cooperative Information Systems, OTM Conferences OTM 2008, Part I*, volume 5331 of *Lecture Notes in Computer Science*, Monterrey, Mexico, 2008. Springer.

[40] Gérald Oster, Hala Skaf Molli, Pascal Molli, and Hala Naja Jazzar. Supporting Collaborative Writing of XML Documents. In *International Conference on Enterprise Information Systems - ICEIS 2007*, Funchal, Madeira, Portugal, 2007.

[41] Charbel Rahhal, Hala Skaf Molli, Pascal Molli, and Stéphane Weiss. Multi-synchronous Collaborative Semantic Wikis. In Gottfried Vossen, Darrell D.E. Long, and Jeffrey XuYu, editors, *10th International Conference on Web Information Systems Engineering - WISE 2009 Web Information Systems Engineering - WISE 2009 10th International Conference Poznan, Poland, October 5-7, 2009 Proceedings.*, volume 5802 of *Lecture Notes in Computer Science*, Poznan, Poland, 2009. Springer.

[42] Hala Skaf Molli, Gérôme Canals, and Pascal Molli. DSMW: a distributed infrastructure for the cooperative edition of semantic wiki documents. In Apostolos Antonacopoulos, Michael Gormish, and Rolf Ingold, editors, *ACM Symposium on Document Engineering (DocEng 2010) Proceedings of the 10th ACM Symposium on Document Engineering*, Manchester, United Kingdom, 2010. ACM.

[43] Hala Skaf Molli, Charbel Rahhal, and Pascal Molli. Peer-to-peer Semantic Wikis. In Sourav S. Bhowmick, Josef Küng, and RolandWagner, editors, *20th International Conference on Database and Expert Systems Applications - DEXA 2009 Database and Expert Systems Applications 20th International Conference, DEXA 2009*, volume 5690 of *Lecture Notes in Computer Science*, Linz, Austria, 2009. Springer-Verlag.

[44] Stavroula Papadopoulou, Claudia-Lavinia Ignat, Gérald Oster, and Moira C. Norrie. Intra/Inter-document Change Awareness for Co-authoring of Web Sites. In James Bailey, David Maier, Klaus-Dieter Schewe, Bernhard Thalheim, and Xiaoyang SeanWang, editors, *International Conference on Web Information Systems Engineering - WISE 2008 Web Information Systems Engineering - WISE 2008*, volume 5175 of *Lecture Notes in Computer Science*, Auckland, New Zealand, 2008. Springer Berlin / Heidelberg.

[45] Mounir Tlili, William Kokou Dedzoe, Esther Pacitti, Patrick Valduriez, Reza Akbarinia, Pascal Molli, Gérôme Canals, and Stéphane Laurière. P2P Logging and Timestamping for Reconciliation. In *34th International Conference on Very Large Databases - VLDB 2008 Proceedings of the VLDB Endowment*, volume 1, Auckland, New Zealand, 2008. ACM.

[46] Mounir Tlili, William Kokou Dedzoe, Esther Pacitti, Patrick Valduriez, Reza Akbarinia, Ludovic Dubost, Sergiu Dumitriu, Stéphane Laurière, Gérôme Canals, Pascal Molli, and Maire Julien. Estampillage et Journalisation P2P pour XWiki. In *8ème Conférence Internationale sur les NOuvelles TEchnologies de la REpartition - NOTERE 2008*, Lyon, France, 2008. ACM.

[47] Mounir Tlili, William Kokou Dedzoe, Esther Pacitti, Patrick Valduriez, Reza Akbarinia, Pascal Molli, Gérôme Canals, Julien Maire, and Stéphane Laurière. Data Reconciliation in P2P Collaborative Applications. In *24ème Journées "Bases de Données Avancées" - BDA 2008*, Guilherand-Granges, France, 2008.

[48] Fodé Touré, Karim Baïna, and Khalid Benali. An efficient algorithm for workflow graph structural verification. In *Coopis 2008 (16th International Conference on cooperative information systems, On the Move*, Monterrey, Mexico, 2008.

[49] Hien Thi Thu Truong, Claudia-Lavinia Ignat, and Pascal Molli. A CONTRACT-EXTENDED PUSH-PULL-CLONE MODEL FOR MULTI-SYNCHRONOUS COLLABORATION. *International Journal of Cooperative Information Systems*, 21(3):221–262, 2012.

[50] Hien Thi Thu Truong, Claudia-Lavinia Ignat, and Pascal Molli. Authenticating Operation-based History in Collaborative Systems. In *ACM 2012 International Conference on Support Group Work*, pages 131–140, Sanibel Island, FL, États-Unis, 2012.

[51] Hien Thi Thu Truong, Claudia-Lavinia Ignat, and Pascal Molli. Securing Logs in Operation-based Collaborative Editing. In *CSCW'12 - The Twelfth International Workshop on Collaborative Editing Systems 2012*, Seattle, États-Unis, 2012.

[52] Stéphane Weiss. *Edition collaborative massive sur réseaux Pair-à-Pair*. PhD thesis, Université Henri Poincaré - Nancy I, 10 2010.

[53] Stéphane Weiss, Pascal Urso, and Pascal Molli. Wooki: A P2P Wiki-Based Collaborative Writing Tool. In Boualem Benatallah, Fabio Casati, Dimitrios Georgakopoulos, Claudio Bartolini, Wasim Sadiq, and Claude Godart, editors, *8th International Conference on Web Information Systems Engineering – WISE 2007 Web Information Systems Engineering – WISE 2007 8th International Conference on Web Information Systems Engineering Nancy, France, December 3-7, 2007 Proceedings*, volume 4831 of *Lecture Notes in Computer Science*, Nancy, France, 2007. Springer Berlin / Heidelberg.

[54] Stéphane Weiss, Pascal Urso, and Pascal Molli. An Undo Framework for P2P Collaborative Editing. In Elisa Bertino and James B.D. Joshi, editors, *4th International Conference on Collaborative Computing : Networking, Applications and Worksharing - CollaborateCom 2008 Collaborative Computing: Networking, Applications andWorksharing 4th International Conference, CollaborateCom 2008 Orlando, FL, USA, November 13-16, 2008 Revised Selected Papers*, volume 10 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Orlando, FL, USA, 2008. Springer Berlin Heidelberg.

[55] Stéphane Weiss, Pascal Urso, and Pascal Molli. Logoot: A Scalable Optimistic Replication Algorithm for Collaborative Editing on P2P Networks. In *29th IEEE International Conference on Distributed Computing Systems - ICDCS 2009 2009 29th IEEE International Conference on Distributed Computing Systems*, Montreal, Canada, 2009. IEEE.

[56] Stéphane Weiss, Pascal Urso, and Pascal Molli. Logoot-Undo: Distributed Collaborative Editing System on P2P Networks. *IEEE Transactions on Parallel and Distributed Systems*, 21(8), 2010.

[57] Ustun Yildiz and Claude Godart. Design and Implementation of Information Flow-sensitive Business Processes. In *IEEE 6th European Conference on Web Services - ECOWS 2008 on Web Services, 2008. ECOWS '08. IEEE Sixth European Conference*, Dublin, Ireland, 2008. IEEE.

[58] Ehtesham Zahoor, Olivier Perrin, and Claude Godart. An Integrated Declarative Approach to Web Services Composition and Monitoring. In Gottfried Vossen, Darrell D.E. Long, and Jeffrey XuYu, editors, *10th International Conference on Web Information Systems Engineering - WISE 2009 Web Information Systems Engineering - WISE 2009*, volume 5802 of *Lecture Notes in Computer Science*, Poznan, Poland, 2009. Springer-Verlag.

[59] Ehtesham Zahoor, Olivier Perrin, and Claude Godart. Rule-Based Semi Automatic Web Services Composition. In *2009 IEEE Congress on Services I - SERVICES 2009 2009 Congress on Services - I*, Los Angeles, CA, USA, 2009. IEEE.

[60] Ehtesham Zahoor, Olivier Perrin, and Claude Godart. DISC: A declarative framework for self-healing Web services composition. In *8th IEEE International Conference on Web Services - ICWS 2010*, Miami, FL, USA, 2010. IEEE.

[61] Ehtesham Zahoor, Olivier Perrin, and Claude Godart. DISC-SET: Handling temporal and security aspects in the Web services composition. In *The 8th IEEE European Conference on Web Services - ECOWS 2010*, Ayia Napa, Cyprus, 2010.

4 RESEARCH PROGRAM

The landscape of computing has greatly evolved in the last 5 years. The vision that we advocate with others [3,7] for ten years regarding the construction of applications through composition of services is happening. Collaborative editing is mainstream. It has a different twist due to important scientific and technological evolutions, mostly because of the "as a service" movement that requires reconsideration of many software engineering techniques. But services are used at an unprecedented scale and collaboration occurs increasingly often, between users and systems, at a scale that is currently not supported by existing tools, since it was not anticipated [2]. We need to refocus and to continue our efforts in order to accommodate this evolution. Our goal is to provide a better understanding and support for high added value data service engineering and provisioning in the age of large scale machine and human computing. We want to help software engineers to make educated decisions when deciding to set up service based software. Contrary to other work in the domain of services, we want to concentrate on a specific family of services,

2. Only 50 users can cooperate on a GoogleDocs (See https://support.google.com/drive/answer/2494827).

[3] Gustavo Alonso, Fabio Casati, Harumi Kuno, and Vijay Machiraju. Web services. In *Web Services*, pages 123–149. Springer Berlin Heidelberg, 2004.

[7] Fabio Casati, Ski Ilnicki, LiJie Jin, Vasudev Krishnamoorthy, and Ming-Chien Shan. Adaptive and dynamic service composition in eflow. In *Advanced Information Systems Engineering*, pages 13–31. Springer Berlin Heidelberg, 2000.

that we call data-centred, starting with services that allows sharing and synchronization of data. Trying to produce results that could be purportedly applied to all kind of services caused a great confusion in the field between people with a Telecom, network, IT or business vision. One of the contributions of our work is that we continually consider problems that occur at the border between the business perspective and the IT perspective and their relationship (Figure 2). These services are very popular but are also very critical since people rely on them for their day to day work. Services like Dropbox [3], Evernote [4] are used by millions of people with very little guarantees. They rely on synchronisation algorithms that are vulnerable to failure. Two years ago, legitimate users of Megaupload lost all their data after a sudden shutdown. Today these services are composed with other services as new services or mashups, sometimes in a transparent way, leaving the users unaware of where their data are. And this is just the beginning. Our expertise at the different levels of the stack for data management services allow us to have a holistic view in this specific area, and a very good understanding of the research problems.

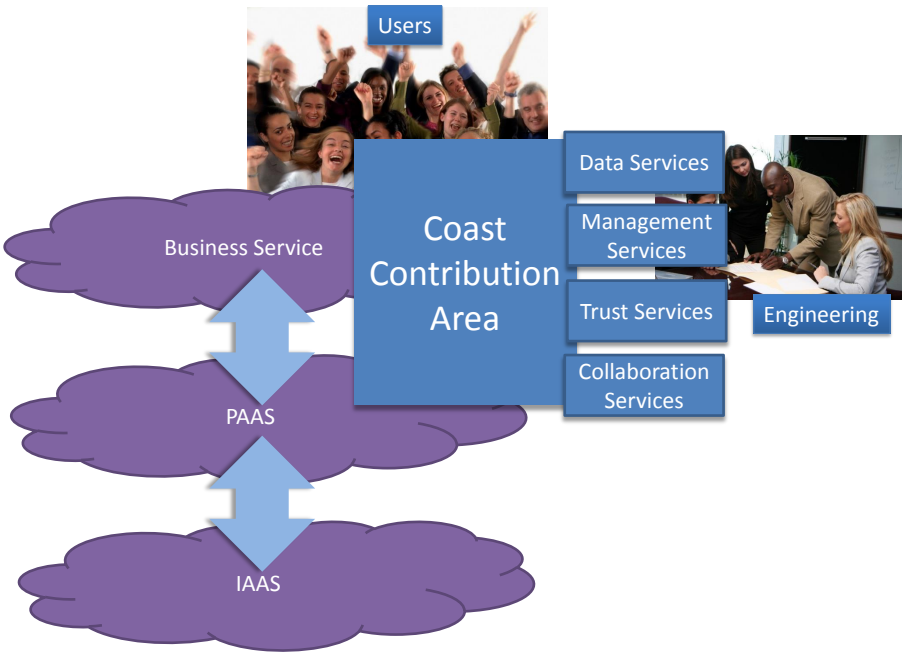

Figure 2: Coast Contribution Area

In order to tackle them effectively, we divide our effort among three sub-projects: large scale collaborative data management, data centred service composition and trustworthy collaborative software systems, the last one being the most challenging and long term one. It relies heavily on the outcome of the first ones.

3. http://www.dropbox.com/
4. http://www.evernote.com/

4.1 LARGE SCALE COLLABORATIVE DATA MANAGEMENT

PARTICIPANTS: Gérome Canals, Claudia Ignat, Gérald Oster, Pascal Urso

Data management services help users and programs to collaborate both synchronously and asynchronously. The two important data management services to be explored are: data sharing, using optimistic data replication to support the merging of concurrent changes on the shared data and conflict avoidance and resolution using awareness mechanisms. In addition, user studies must examine requirements of these services and validate the proposed solutions.

Collaborative editing requires data sharing. In order to provide efficient data availability, data is typically replicated. Optimistic replication algorithms provide a proven, effective approach to supporting the data sharing dimension of collaboration under limited conditions. The behaviour of optimistic algorithms is poorly understood under intensive use in different architectural settings, with different scenarios and different kind of data structures. Previously well-established optimistic replication approaches such as operational transformation work well on groupware settings with few number of users, but do not adapt well to large scale settings with a large number of users making frequent modifications on the shared documents. Furthermore, several existing solutions impose constraints either on the system architecture such as client/server-based architectures [19] or on the processing order of generated modifications [25] both of which limit extension to large scale settings. Solutions that do not impose these constraints use data structures that are a function of the number of users such as state vectors [11,24] or/and of the number of operations such as history buffers [11,24,5]. Hence, the time complexity of existing algorithms is proportional to the total number of users and to the total number of operations. Both of these threaten scalability to large settings. Alternatively, CRDTs provide a more recent optimistic replication solution with weaker requirements, whose underlying data structures are independent of the total number of users and operations. However, CRDTs were developed only for linear structures of fixed granularity and they suffer from meta-data overhead.

One of the key challenges is to provide optimistic replication solutions that support a large number of users with a high frequency of concurrent mutations while respecting user expectation and minimizing meta-data overhead. New algorithms for new data types are required, possibly combining pessimistic and optimistic approaches. These algorithms should accomodate different granularities of data for different types of complex communication data in addition to text such as wikis, key-value map structures (used by DHTs [22], Facebook's

---

[19] David A. Nichols, Pavel Curtis, Michael Dixon, and John Lamping. High-latency, low-bandwidth windowing in the Jupiter collaboration system. In *Proceedings of the 8th annual ACM symposium on User interface and software technology - UIST '95*, pages 111–120. ACM Press, 1995.

[25] Nicolas Vidot, Michèle Cart, Jean Ferrié, and Maher Suleiman. Copies Convergence in a Distributed Real-Time Collaborative Environment. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work - CSCW 2000*, pages 171–180, Philadelphia, Pennsylvania, USA, December 2000. ACM Press.

[11] Clarence A. Ellis and Simon J. Gibbs. Concurrency Control in Groupware Systems. *SIGMOD Record: Proceedings of the ACM SIGMOD Conference on the Management of Data*, 18(2):399–407, May 1989.

[24] Chengzheng Sun, Xiaohua Jia, Yanchun Zhang, Yun Yang, and David Chen. Achieving Convergence, Causality Preservation, and Intention Preservation in Real-Time Cooperative Editing Systems. *ACM Transactions on Computer-Human Interaction*, 5(1):63–108, March 1998.

[5] Michelle Cart and Jean Ferrié. Asynchronous reconciliation based on operational transformation for P2P collaborative environments. In *Proceedings of the International Conference on Collaborative Computing: Networking, Applications and Worksharing - CollaborateCom 2007*, pages 127–138, White Plains, New York, USA, November 2007. IEEE Computer Society.

[22] Rodrigo Rodrigues and Peter Druschel. Peer-to-peer systems. *Communications of the ACM*, 53(10):72–82, October

Cassandra [17], Google's BigTable [8]) and semi-structured documents with invariants. We believe promising solutions for minimizing meta-data overhead hinge on exploitation of the underlying data structures [4]. Any proposed algorithm requires both theoretical and practical evaluation in terms of time and space complexity [1]. Evaluation will include analysis of histories of human or machine produced data modifications (e.g. in Google Docs or distributed version control systems such as Git, Mercurial) in the context of the cloud environment and peer-to-peer networks. We have collaborations with data producers such as Rovio (http://www.rovio.com/) and Basho (http://basho.com/) in the context of the European Syncfree project and XWiki (http://www.xwiki.org/) in the context of ANR XWiki Concerto and STREAMS. We plan to perform evaluations on the data obtained from these data generators and other ones such as Google.

Second, conflicting modifications arise when people concurrently modify a shared document. In the period of time between the occurrence and discovery of the conflict, it may grow and become difficult to resolve. Resolving these conflicts is time consuming and error prone, as it requires a precise identification of the concurrent changes and the context in which they were performed [12,21]. Web-scale collaboration exacerbates the problem. Conflicts occur more often, involving a larger group of users for resolution. A potential solution lies in prevention via timely warning, so that users are aware of a potential problem. However, existing awareness mechanisms in collaborative editing targeted only groupware with few users. Moreover, once conflicts occur, the challenge lies in identifying, representing, visualizing and classifying them in a manner that helps resolution at a large scale. We will investigate solutions that anticipate the notion of conflict within the underlying optimistic replication algorithms.

Finally, the constraints of different collaboration modes and how users face them have never been studied. Furthermore, users can generate different real world use cases that pose different resource intensive challenges for optimistic replication mechanisms and awareness solutions. We plan to perform empirical studies with users to understand the computational requirements of different collaboration scenarios and whether users adapt to the constraints of existing systems. This has a potentially crucial impact on the underlying optimistic replication algorithms, which may require optimization to attain user acceptance. We will employ smaller scale studies to determine awareness mechanism requirements and larger scale studies to test proposed awareness protocols and settings. System evaluation from a user's point of view can employ several research traditions, including activity theory and cognitive engineering, both of which are readily available in Europe. However, expertise in experimentally-oriented user studies centers in the United States, where studies in collaborative work represent the cutting

---

2010.

[17] Avinash Lakshman and Prashant Malik. Cassandra: a decentralized structured storage system. *ACM SIGOPS Operating Systems Review*, 44(2):35–40, April 2010.

[8] Fay Chang, Jeffrey Dean, Sanjay Ghemawat, Wilson C. Hsieh, Deborah A. Wallach, Mike Burrows, Tushar Chandra, Andrew Fikes, and Robert E. Gruber. Bigtable: a distributed storage system for structured data. In *Proceedings of the 7th USENIX Symposium on Operating Systems Design and Implementation*, OSDI 2006, pages 15–15, Berkeley, CA, USA, 2006. USENIX Association.

[4] Luc André, Stéphane Martin, Gérald Oster, and Claudia-Lavinia Ignat. Supporting Adaptable Granularity of Changes for Massive-scale Collaborative Editing. In *Proceedings of the International Conference on Collaborative Computing: Networking, Applications and Worksharing - CollaborateCom 2013*, Austin, Texas, USA, October 2013.

[1] Mehdi Ahmed-Nacer, Claudia-Lavinia Ignat, Gérald Oster, Hyun-Gul Roh, and Pascal Urso. Evaluating CRDTs for Real-time Document Editing. In *Proceedings of the 11th ACM Symposium on Document Engineering*, DocEng 2011, pages 103–112, Mountain View, California, USA, September 2011. ACM Press.

[12] Jacky Estublier and Sergio Garcia. Process model and awareness in scm. In *Proceedings of the 12th international workshop on Software configuration management*, SCM 2005, pages 59–74, New York, NY, USA, 2005. ACM.

[21] Dewayne E. Perry, Harvey P. Siy, and Lawrence G. Votta. Parallel changes in large-scale software development: an observational case study. *ACM Transactions on Software Engineering and Methodology*, 10(3):308–337, July 2001.

edge of usability studies. To provide this crucial expertise, we collaborate with the Department of Psychology of Wright State University via the USCoast associated team established in 2013.

We will pursue several main research directions in pursuit of large scale collaborative data management.

- *New optimistic replication algorithms for new data types*
- *Theoretical evaluation of these algorithms*
- *Practical evaluation of these algorithms under different scenario and with different scales*
- *Group awareness mechanisms for communities in different contexts*
- *Experimental studies on requirements of replication algorithms and on their evaluation*
- *Evaluation of group awareness mechanisms by means of user studies*

Some work is now ongoing on these topics regarding design and evaluation by means of simulations of new algorithms [6, 3]. We proposed a framework [2] to automatically evaluate the effort required for users to use a given merge algorithm using a large scale corpus of collaborative editing. We use this framework to study and improve the impact of the different families of algorithms on the result of the merge operations [1]. We performed some experimental user studies to study real-time constraints in collaborative editing [35]. We plan to construct a new collaborative editing platform that exploits the benefits of the different algorithms in order to conduct controlled experiments and to deploy it on a large scale.

4.2 DATA CENTERED SERVICE COMPOSITION

PARTICIPANTS: Khalid Benali, François Charoy, Claude Godart, Olivier Perrin, Samir Youcef

Service Oriented Computing is now a recognized research field in computer science originating from different areas like component based software engineering and web service standards [20]. Among the research directions that researchers have followed in this field, the one that tries to understand the mechanisms of service assembly and composition is probably one of the most prolific. Many techniques have been explored to automate or verify composition and orchestration. Concurrently, we have witnessed the advent of the Cloud that provides an elastic infrastructure that supports a growing number of public web-based API [5]. This is a pragmatic outcome of the original vision of Service Oriented Architecture and service orchestration with a business perspective [2]. Companies and government agencies not only provide their data or business services through a Web application but also through an API that can be used to create third party applications or composed with other services. For example, in March 2013, there were 183 services available in the travel directory of ProgrammableWeb, 95 of which were based on the REST protocol and 50 on the SOAP protocol. From our point of view, this eliminates the idea of automatic composition, as recognized recently by Fabio Casati [6].

Today, developers face difficult problems in combining these heterogeneous services, provided

5. See http://www.programmableweb.com/

[20] Michael P. Papazoglou, Paolo Traverso, Schahram Dustdar, and Frank Leymann. Service-oriented computing: State of the art and research challenges. *Computer*, 40(11):38–45, 2007.

[2] Gustavo Alonso and Fabio Casati. Web services and service-oriented architectures. In *Proceedings of the 21st International Conference on Data Engineering*, ICDE 2005, pages 1147–1147, Washington, DC, USA, 2005. IEEE Computer Society.

[6] Fabio Casati. Promises and failures of research in dynamic service composition. In Janis Bubenko, John Krogstie, Oscar Pastor, Barbara Pernici, Colette Rolland, and Arne Sølvberg, editors, *Seminal Contributions to Information Systems Engineering*, pages 235–239. Springer Berlin Heidelberg, 2013.

by different companies, with very different terms of use. For instance, a very common term of use limits the number of function calls per period of time. Others limit the bandwidth use, or the number of concurrent user for a given customer. Developers also face the problem of managing the different service level agreements that can be negotiated with each service [16]. When trying to compose services, developers must consider both the functional dimensions of services and their non-functional ones, including these properties that cannot be controlled by the client, like service business models, service capacity and temporal constraints. Today, the service provider describes these constraints in terms of service contracts as plain text. Services are described as very basic interfaces. The difficulty in evaluating the feasibility and risk of service composition limits its potential. The service composition problem appears at design time but also at runtime and service modification makes it more complex. From a software engineering perspective, we face a recurring challenge regarding software construction : how to build software that can be relied upon, using service components whose operation is controlled by a third party. As a result

- the service capacity is not controlled
- the service capacity is shared among different unknown clients
- the service operation can be stopped by the provider
- the service API can change
- the service behaviour can change with our without notice

Our aim is to design and validate models, methods and tools to engineer, deploy, manage and compose services safely and efficiently. This research focuses on the engineering of reusable software. The question of software quality extends to the availability of the software in question on an IT infrastructure for composition and reuse. The business model of the service is also potentially important. We must investigate these issues in relation to the functions that the runtime infrastructure delivers. We will focus on services that rely on the data sharing infrastructure work described above. We must provide developers with a precise description of the Terms of Service and of the Service Level Agreements for different kinds of business plans for the consumption of these services. For data sharing it can be related to the size of the data stored, the bandwidth used to access it, the number of read or write operations, or the level of availability required. Then we must support developers in selecting and combining services based on these dynamic attributes and to check the validity of this composition or to make it achievable by using the right service at the right time [26]. It will be possible to derive new terms of service and to infer the quality of service that we may expect. We must provide a reliable framework that supports business analyst and software engineer in the design, the deployment, the maintenance and the composition of web based services that exhibit enforceable non-functional properties.

Our approach here will continue our current work, to propose models and tools that allow enforcing and validating properties of service composition, at both design time [10,28] as well as runtime [27].

---

[16] Kyriakos Kritikos, Barbara Pernici, Pierluigi Plebani, Cinzia Cappiello, Marco Comuzzi, Salima Benrernou, Ivona Brandic, Attila Kertész, Michael Parkin, and Manuel Carro. A survey on service quality description. *ACM Comput. Surv.*, 46(1):1:1–1:58, July 2013.

[26] Jun Yan, Ryszard Kowalczyk, Jian Lin, Mohan B. Chhetri, Suk Keong Goh, and Jianying Zhang. Autonomous service level agreement negotiation for service composition provision. *Future Generation Computer Systems*, 23(6):748 – 759, 2007.

[10] Karim Dahman, Francois Charoy, and Claude Godart. Towards consistency management for a business-driven development of soa. In *Proceedings of the 2011 IEEE 15th International Enterprise Distributed Object Computing Conference*, EDOC 2011, pages 267–275, Washington, DC, USA, 2011. IEEE Computer Society.

[28] Ehtesham Zahoor, Olivier Perrin, and Claude Godart. Disc-set: Handling temporal and security aspects in the web services composition. In *Proceedings of the 2010 Eighth IEEE European Conference on Web Services*, ECOWS 2010, pages 51–58, Washington, DC, USA, 2010. IEEE Computer Society.

[27] Ehtesham Zahoor and d Claude Godart Olivier Perrin a. An event-based reasoning approach to web services monitoring. In *Proceedings of the 2011 IEEE International Conference on Web Services*, ICWS 2011, pages 628–635,

- *We will define models for the formal description of Term of Services and Service Level Agreement for different family of services, recognizing that one size does not fit all*
- *We will study tools and methods to validate and certify compositions. Depending on the nature of services, different kinds of approaches may have to be selected. Today, we are considering event based approach to describe and reason about composition but a prescriptive model based on orchestration models and on patterns may be more suitable in some cases.*
- *We will define and validate methods for service and service composition engineering, taking into account the economical dimension of service operation. API and the cloud make the economical dimension of service composition explicit, associating a monetary cost to service consumption based on calls, bandwidth, storage or other attributes.*

We will focus on the family of services that we call data centred, recognizing that all kind of services cannot be treated in the same way. However, we do not limit ourselves to this family and we will explore to what extend the work that we are doing can be adapted to other family of services.

This work is currently ongoing, on service for business process management in the Cloud [12, 13]. We are also working on the monitoring of choreographies in order to detect problems during their execution [7, 9].

4.3 TRUSTWORTHY COLLABORATIVE SOFTWARE SYSTEMS

PARTICIPANTS: Khalid Benali, François Charoy, Claude Godart, Claudia Ignat, Gérald Oster, Olivier Perrin, Pascal Urso, Samir Youcef

A collaborative software system is a system of associate software services and users to serve a defined business need. Let us consider the example of a user using a composition of two services, Google Agenda and Doodle. He wants to allow Doodle to consult Google Agenda in order to provide his availability for meetings scheduled with Doodle and to update it in the case a meeting is scheduled. First of all note that neither Doodle nor Google Agenda provide guarantees for the services they offer, such as availability, fault tolerance, provided performances or data persistency. Specifying the guarantees for the composition of these two services is even more complex. For instance what is the guaranteed response time for an update done in Google Agenda to be visible in Doodle? What would be the acceptable delay for users ? Moreover, Doodle needs detailed access rights to read and write to Google Agenda. For instance, when consulting Google Agenda, Doodle should get access only to availabilities for the specified time periods and no information about the activities performed. Specifying such detailed access rights is tedious and error-prone, particularly at a large scale. This is one of the problems that we are investigating in the OpenPaaS project.

The potential for combined services, along with the number of users, and the number of use cases create novel computer science challenges. For example, crowdsourcing tasks, such as voluntary contributions to Wikipedia, or paid workers using Mechanical Turk, are rapidly growing in scope and ambition [15]. When services are composed, no guarantee exists that they will function together in a manner that preserves user expectations concerning service quality, performance, the availability synchronization and data persistence. User trust is the foundation of service selection. In fact, what is actually happening in our Doodle example is that Google Agenda

Washington, DC, USA, July 2011. IEEE Computer Society.

[15] Aniket Kittur, Jeffrey V. Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. The future of crowd work. In *Proceedings of the 2013 conference on Computer supported cooperative work*, CSCW '13, pages 1301–1318, New York, NY, USA, 2013. ACM.

and Doodle have full rights to read and modify the other application. However, people still use these services because they trust them, while ignoring the guarantees they offer. Users require assurance of reliable service behaviour, and feedback when the service does not behave as expected.

We envision a sophisticated infrastructure which leverages the other key focuses of this work in collaborative data management and service composition. The infrastructure will enable the development and deployment of services that will allow users or computers to make educated decisions, balancing risk and expectation while supporting a reliable collaboration among communities of users.

We require a comprehensive approach to the question of confidence, trust and of trustworthiness for a system of associate software services. Here, we define trust and the related constructs, confidence and control consistent with [9]. Trust reduces transaction complexity between different partners by avoiding the need for costly and time consuming control. When a system becomes too complex, when the number of stakeholders is very high, it becomes prohibitively expensive if not impossible to control every condition regarding the expected outcome of a given situation. Thus control and trust trade off in the decision to use a system or to consume a service. Our capacity to control a system and/or trust it along with risk mitigation possibilities influences the risk we assume regarding an outcome. We consider reliance on trust to be essential for a large scale orchestration of services and communities of users. This demands a multi-faceted approach that spans both technical and organisational considerations.

From the perspective of service availability, trust depends upon different attributes related to the provider of the service, the way the service is delivered, the purpose of this service consumption and its cost. How does this trust accrue and evolve as a part of a business reputation? Today such concerns remain uninvestigated for services such as data sharing or processes that require a combination of user actions and service invocations. We need a framework or model to provide a trustworthy infrastructure for service composition including the constraints and principles that govern the processes and their compositions.

Regarding service composition, third party services are already accessible via Web protocols. However, the external nature of these services challenges the maintenance of data integrity and privacy, data and operational availability, accountability and compliance. Current certification schemes are insufficient. We cannot use them to support and automate run-time security assessment such as access control.

As mentioned above, specifying such detailed policy is tedious and error-prone, particularly at a large scale. Our envisaged solution is to give access without control and check a-posteriori if the guarantees were respected. Respect or violation of the guarantees would trigger the update of the trust value in these services.

We will provide a new framework to evaluate trust in a service, based on its attributes, such as quality, manner of provision, and the risk of failure from the perspective of the stakeholders. A vendor/suppliers' contract can predefine terms of service or some other more formalized service level agreement. Among users, the agreement may function as predefined rules about updates or distribution of data. However, in order to monitor system functionality and provide feedback to users, we require a model that encompasses the attributes that define trustworthy collaboration.

The guidelines described by Schneiderman [23] show that trustworthiness is a very multidimensional and

[9] Piotr Cofta. *Trust, Complexity and Control - Confidence in a Convergent World*. Wiley, 2007.

[23] Ben Shneiderman. Designing trust into online experiences. *Communications of the ACM*, 43(12):57–59, December

requires careful validation. More precisely:

- *we will define models for contracts between services based on collaboration and service composition. These models enable description of the promises made to service users but also service constraints and costs.*
- *we will identify confidence attributes for data sharing, collaboration and services. These values depend not only on technical factors but also on the usage scenario.*
- *we will provide support to verify trust contract during or after the execution of services or composition of service. This is an important factor for the delivering of a trustworthy system. Given the scale we consider this could be done after execution, in real time or a posteriori for large scale combination of service and human. Again, we believe that there is no unique solution and that different kind of services or contexts will require different approaches.*

This collaboration model requires validation in web scale collaboration settings. We must also understand how to measure and then encourage user trust in service availability, service performance, data quality, data availability, as well as other users for the successful completion of a transaction. Reputation models [18] will be support for evaluation in our specific context. This is something that we are currently studying [49, 51, 50].

We must conduct a reassessment of security and risk, and propose new schemes and techniques fully integrated within the SOA lifecycle for expressing, assessing and certifying security and safety properties. This is also currently ongoing in the context of Cloud based architecture [26, 25]

We plan to conduct user studies for the evaluation of the proposed model. In order to validate it at a large scale we propose using a crowdsourcing scenario illustrating the composition of services described together with their constraints. We have started some work on this topic as well [11, 10].

4.4 METHOD

In the area that we are exploring, it is difficult to produce useful results following either a purely experimental or a theoretical approach. Understanding the relation between our technical artefacts, users and organizations, requires a method for the investigation, design, development and evaluation of the service that we want to assess. We rely on the general theory of Design Science [14] as a foundation for our research method.

The system and the protocol we want to study and experiment with do not exist. They must be developed, based on theories that we are producing or on competing ones, to enable usage and evaluation. We are also facing the recurring question of the scale of our experimentations. Today, we have several directions that we need to investigate, such as live experimentation with many users, the use of a large scale collaboration corpus (e.g. Wikipedia or open source version control logs) or the production of our own corpus, and experimentation in changing contexts that occurs thanks to the increased user mobility. It must be noted that very little rigorous work addresses the evaluation and comparison of distributed collaborative settings. Thus, as it is the case today,

---

2000.

[18] Sergio Marti and Hector Garcia-Molina. Taxonomy of trust: categorizing p2p reputation systems. *Computer Networks*, 50(4):472–484, March 2006.

[14] Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram. Design science in information systems research. *MIS Quarterly*, 28(1):75–105, March 2004.

we are adapting our evaluation method to the system we want to evaluate and to the conditions of evaluation. We have used simulation to compare the behaviour of replication algorithms and to evaluate crowdsourcing patterns, experimentation to characterize "real time" collaborative editing, and we plan to use ethnography to evaluate crowdsourcing in cities. We will continue to adapt our method to the context and the hypotheses we want to test. Combining crowdsourcing and collaborative editing could be a mean to evaluate our platforms and algorithm at a large scale in a repeatable way and in different context. Mobility is also an important dimension. People are collaborating on the move and are consuming services from heterogeneous mobile devices. This opens new challenges for our work, especially regarding the way to evaluate our results in all these conditions and to measure the impact of context on collaboration and service consumption.

This approach that consist in the development of software for experimentation and evaluation is a trademark of our team and has led to some transfer. Indeed Bonita and LibreSource initially arose from experimentation purposes.

5 APPLICATION DOMAIN OR MOTIVATIONAL STORIES

Our research will have applications in several important domains such as software engineering, business process management and large scale collaboration in general. Moreover it will leverage existing platforms to provide advanced data sharing services and service supports while exhibiting measurable trust attributes.

5.1 CLOUD BASED FORGE FOR SERVICE BASED APPLICATION DEVELOPMENT

Application development, testing and deployment in a Cloud based environment burdens developers. The heterogeneity of services, of libraries, of devices and of environments makes it very difficult for teams to synchronize and control their development. An emerging solution is to propose a development platform fully based on a Cloud infrastructure that would allow a team of developers to share their resources, to develop, test and deploy their applications in a shared environment available through any kind of targeted device. The contemporary state of the art relies on the assembly of old and unproven technology for data sharing and service composition [6]. Thus our work, in its different dimensions, and mostly the long term one, should allow us to deliver a trustworthy cloud based software forge built upon reliable data sharing service together with a composition of the required and trustworthy services for software development [7]

The research that we aim to do could be applied to provide advanced and reliable software development services covering the whole development and production cycle. The forge federating application unifies the challenges discussed previously:

6. Cloud9 and Codenvy are cloud based development environments that tolerate development and deployment software in the Cloud. Still they rely on a traditional version management system and leave the developer very seldom if he wants to compose services proposed by the platform.

7. These services are legion today.

- A forge presents a composition of very heterogeneous services based on a data sharing service as we have demonstrated it in the FP6 QualiPSo project.
- A forge can procure services to support the composition and the deployment of external services. This is done in a very ad-hoc way in the first available cloud based development environments.
- A forge can provide services for the verification, composition, test, deployment and management of service based applications.
- A forge is by nature a collaborative environment where data of different kind are produced and shared with different policies and potentially among different organizations.
- Adoption of a forge and of its services by a community of developers requires that these developers trust the forge to deliver on its promises and that customers of the services trust these services and their composition.
- Software development is now conducted by communities that can be very large, combining a wide scope of diverse expertises and skills.

The forge scale provides an interesting test bed for our work on collaborative data management since a forge can be used to host and deploy thousands of projects and since hundreds of users may contribute to a project.

Then, forges are also a very good test bed to validate trust and reputation models, both for their use as a system and for their ability to deliver trusted services. Furthermore, the social network of developer and project manager can support sophisticated trust models. For us, a software forge is an ideal application regarding all the dimensions that we want to explore. Our service based approach should allow us to transfer some of our result more quickly if they can be delivered by existing environment rather than by trying to build once again an entire environment of our own.

5.2 LARGE SCALE MULTI-ORGANIZATIONAL PROJECT MANAGEMENT (A GENERALIZATION OF A FORGE)

Increasingly often, resolving the problems faced by humanity requires large scale and international collaboration. Prominent examples are the writing of the Climate Change Report [8] or more recently the initiative that has started to assess the state of Biodiversity [9] in the world. More dramatically, disasters occur randomly in different places in the world and may require the coordination of many organizations from different countries to help the populations involved [13]. Supporting cooperation and coordination at this scale, while providing guarantees regarding the privacy and the integrity of the collected data, is still an issue. Recent advances rely on the good will of large corporations that release their resources [10]. At a more manageable scale, the initiatives around the transformation of cities to answer the challenges of the XXIst century requires the leverage of citizen participation. This also requires coordination at a large scale that

8. See http://www.ipcc.ch/ (Climate Change)
9. See http://ipbes.net/ (Biodiversity)
10. See http://www.google.org/crisisresponse/ (Google and Crisis Management)

[13] Jörn Franke, v Charoy, and Cédric Ulmer. Coordination and situational awareness for inter-organizational disaster response. In *10th International IEEE Conference on Technologies for Homeland Security*, HST 2010, Waltham, Boston, MA, États-Unis, November 2010.

we aim to support.

To summarize, people could use our results to set up trustworthy collaboration infrastructures free from political and geographical constraints and based on less capital intensive environments that are data centers.

5.3 SERVICE FOR DIGITAL SERVICE DELIVERY PLATFORM

As a generalization of our work we will consider conducting work on services for digital service delivery. Digital service delivery is becoming a business by itself. We can observe the emergence of start-up in this domain like Mashape (https://www.mashape.com/) or like the addons of public clouds like on Heroku. They provide services to make available digital services like data sharing, communication services or others to users and developers, to control, monitor and monetize their use by applications and other services. This is highly valuable from the service provider point of view. It does not take into account yet the consumer point of view. As we explained it before, these initiatives are very ad-hoc and our research could provide the knowledge and the technology to bring this new way to build software to a new standard. This idea of Service Delivery Platform is not new. It takes its root in the old idea of software component library and its most probably a long term challenge.

6 COLLABORATIONS AND GRANTS

INTERNATIONAL ACTIONS.

- **USCoast Associated Team** (2013–2015), User Studies on Trustworthy Collaborative Systems, http://uscoast.loria.fr/

PAST ACTIONS.

- **VanaWeb Associated Team** (2008–2010), Hybrid and Autonomous Constraint Solving: Composition Problems for The Web, http://www.loria.fr/~ringeiss/CHILI/vanaweb/index.html

EUROPEAN ACTIONS.

- **SyncFree** (FP7-609551, 2013–2016) is a Specific Targeted Research Project on Large-scale Computation Without Synchronisation, http://syncfree.lip6.fr/
- **City CrowdSource** (2013), is an EIT ICT Lab activity belonging to the "Digital Cities of the Future" action line, http://www.eitictlabs.eu/innovation-areas/digital-cities-of-the-future/

PAST ACTIONS.

- **Integrated Project QualiPSo** (FP6-IP-034763, 2006–2011), Trust and Quality in Open Source Systems, http://qualipso.org/
- **Network of Excellence INTEROP** (2004–2007), Interoperability Research for Networked Enterprises Application and Software, http://www.interop-vlab.eu/

NATIONAL ACTIONS.

- **FSN OpenPaaS** (2012–2015), funded by the "Fond National pour la société Numérique", aims at developing a PaaS (Platform as a Service) technology dedicated to enterprise collaborative applications deployed on hybrid clouds (private/public),
  http://www.open-paas.org/
- **ANR Blanc ConcoRDanT** (ANR-10-BLAN-0208, 2010–2014), CRDTs for consistency without concurrency control in cloud and peer-to-peer systems,
  http://concordant.lip6.fr/
- **ANR Arpège STREAMS** (ANR-10-SEGI-010, 2010–2014), Solutions for peer-to-peer real-time social web,
  http://streams.loria.fr/

PAST ACTIONS.

- **Wiki 3.0** (2009–2012), funded by the "Ministère de l'Économie, des Finances et de l'Industrie", a new wiki generation which offers real-time editing and integration of social interaction tools,
  https://wiki30.xwiki.com/
- **Coclico** (2009–2011), funded by the "Pôles de Compétitivité SYSTEM@TIC and MINALOGIC" , COnvergence de la Communauté Libre des Infrastructures Collaboratives Ouvertes,
  http://www.projet-coclico.org/
- **ADT Galaxy** (2008–2010), An Open SOA agile platform, funded by Inria,
  http://galaxy.gforge.inria.fr/
- **iCRISIS-Gestion des crises** (2007–2009), funded by the "Ministère de l'Écologie, du Développement et de l'Aménagement Durables"
- **ANR XWiki Concerto** (2006–2009) aiming at evolving the open-source XWiki engine toward a P2P architecture,
  http://concerto.xwiki.com/
- **INRIA ARC RECALL** (2006–2007), Réplication optimiste pour l'Édition CoLLaborative massive sur réseau pair-à-pair,
  http://recall.loria.fr/

PAST REGIONAL ACTIONS.

- **CyWiki** (2008–2010), multi-disciplinary project funded by the University of Lorraine, a software infrastructure for the collaborative and assisted transformation of textual content into formal and structured knowledge
- **PSW (Proof of Services Web)** (2007–2008), action of the "CPER Modélisation, Information et Systèmes Numériques",
  http://sites.google.com/site/preuvesdeservicesweb/
- **COWS (Contraints for Composition of Web Services)** (2006–2007), action of the "CPER Intelligence Logicielle",
  http://www.loria.fr/~operrin/QSL/

OTHER COLLABORATIONS.

- From 2006 to 2012 we have maintained **a continuous relationship with SAP Research in France, Germany and Australia**. This has resulted in the funding of two PHD thesis, a visiting position for 6 months, several publications and a patent [23, 22].
- Francois Charoy has been invited for 5 months at **NICTA in Sydney** between March and July 2011. A collaboration with **the team of Anna Liu**, Research Group Leader for Software Systems, has been settled during that stay.
- We are working with **Prof. Valerie Shalin of the Department of Psychology from Wright State University** (USA). After a two months visit in 2012 and a two months visit in 2013, we are now collaborating in the framework of the USCoast Inria associate team (http://www.inria.fr/en/teams/uscoast/). She started in October 2013 her sabbatical year with the Score team. This collaboration is strategic for our team concerning experimental and user studies on collaborative methods and systems.
- We are working with **Nuno Preguiça from Universidade Nova de Lisboa** (Portugal) on replicated data types. Pascal Urso and Mehdi Ahmed-Nacer visited his team for several

weeks.
- We collaborate with Prof. Weihai Yu from Department of Computer Science, University of Tromsø on commutative replicated data types. He started in September 2013 his sabbatical year in the Score team.
- We collaborate with Marc Shapiro from REGAL project-team at Inria on commutative replicated data types through the projects ANR ConcoRDanT, ANR STREAMS and the european project SyncFree.
- Olivier Perrin is member of the Bananas (Automated design and autonomous control of hybrid solver cooperations) Inria associate team between Cassis project-team at Inria and UTFSM, Valparaíso, Chile and CMM, Santiago, Chile.

As we can see from the listed on-going collaborations we plan some strategic collaborations with key experts in data consistency from a system point of view. In this area, we collaborate with the main actors in France and Europe. Regarding the service dimension, we need to consolidate our relationships with actors in industry and in research. This is already the case in France, but not as much as we would like in Europe or in the world. We participate in the BPM Round Table that convenes major international actors in the domain of BPM. We are also trying to set up collaboration with team that brings us competencies that we need in the domain of social sciences. This is the case currently on Human Factors with the University of Wright and on Innovation with the ERPI Team of University of Lorraine and the Lorraine Smart City Living Lab. We are preparing the organisation of the ISCRAM Med conference, an international conference that will be held in Toulouse on Information Systems for Crisis Management around the Mediterranee

7 SOFTWARE

- **Replication Benchmarker** (http://github.com/score-team/replication-benchmarker/) is a performance evaluation framework for replication mechanisms used in collaborative editing applications [2]. It contains a library of replicated data types and helpers to create randomized corpus or real corpus extracted from git DVCS histories. This software is under development and is currently used by the members of our team. Software self-assessment: A-3, SO-3, SM-3, EM-2, SDL-4.

PAST ACTIONS.

- **Bonita** is a workflow engine (http://www.bonitasoft.com/). Its development has been initiated in 2001 in the Ecoo team with the contribution of Miguel Valdes, engineer in the team. Bonita was developed as a result of research conducted on collaborative workflow. Since 2004, Bonita has evolved into a real product under the lead of Miguel Valdes, sponsored by Bull. In 2009, Miguel Valdes founded the BonitaSoft Company with two partners to continue its development as an open source software. BonitaSoft has raised more than 5M$ in venture capital. Even though we are not directly contributing to the development of Bonita, this result consolidates this successful transfer from the ECOO team and the work we did at that time.
- **QualiPSo Factory** (http://qualipso.gforge.inria.fr/) is a SOA-based (Service Oriented Architecture) next generation forge developed within the QualiPSo European Project. With the QualiPSo Factory, we have demonstrated that it was possible to build a software forge by composing different kind of services.
- **XWiki Concerto** (http://concerto.xwiki.org/) is a peer-to-peer extension of the open-source enterprise XWiki system developed during ANR XWiki Concerto. It is available in the XWiki repository and interested users can install it.
- **Wooki** (http://wooki.sourceforge.net/) is a P2P wiki system based on WOOT. It has a semantic extension called **Swooki**.
- **DSMW** (http://www.mediawiki.org/wiki/Extension:DSMW) is an extension to Semantic Media Wiki (SMW), the semantic version of MediaWiki. It is publicly available in the official MediaWiki source repository maintained by the MediaWiki community.
- **WikiTaable** (http://taaable.fr/) is a distributed collaborative knowledge building systems based on DSMW written in

collaboration with LIRIS and Inria/Orpailleur.

- **Wiki3.0** is a real-time WYSIWYG editor of wikis. The solution that we proposed in the context of Wiki 3.0 project (https://wiki30.xwiki.com/) is available as an extension to XWiki (http://extensions.xwiki.org/xwiki/bin/view/Extension/RealTime+Wiki+Editor). Software self-assessment: A-4, SO-3, SM-3, EM-3, SDL-4.
- **Rivage** (Real-tIme Vector grAphic Group Editor) (https://github.com/stephanemartin/rivage) is a real-time collaborative graphical editor. Software self-assessment: A-3, SO-3, SM-2, EM-2, SDL-4.

We have a long history of software production, either internally or as part of projects with partners. This software has been used for experimentations and validation. Some have been put into production.

In the future, we plan to continue our software development effort as part of our research method. We will leverage the cloud to distribute our future development as services rather than as just libraries or components. Cloud infrastructure simplifies the operation of software services while keeping the control on the software development process. Maintaining several versions and conduction A/B testing experiments is much more easier. This will obviously help us to validate our research questions.

8 OPPORTUNITIES

8.1 SCIENTIFIC OPPORTUNITIES

The proposed program of research includes an important experimental component. Conducting experimentation is hard and time consuming. Collaborations with external partners facilitate this work. Our experience with the ECOO project-team and the SCORE team is that being an Inria project-team provides visibility, collaboration opportunities and man power (interns, engineers, post-docs).

The research issues we plan to address are not addressed elsewhere within Inria as far as we know and fit with different dimensions of the new strategic plan. The emphasis that we place on trust and on the development of trustworthy collaborative systems traverses three dimensions of the scientific strategy. Trust is a central property to promote the adoption of collaborative systems. In the coming years, in large organizational or community settings, trust models will surely supersede the traditional security model.

We also emphasis in our research and in its applications on the relations between humans and computers, at the border of several disciplines, considering big societal challenges. This resonates very well with the introduction to the new strategic plan since we stand at a very unique place in the Inria research landscape, encompassing the socio-technical vision of the digital society, trying to understand how the loop between users, services and technology operates and how we can make it better and more predictable.

8.2 RELATION WITH INRIA PROJECTS

A preliminary study shows that the proposed project has intersections with several other Inria projects. We stand between very technical projects that consider technical issues at the algorithmic and at the IT level and user centred projects. The former projects address data and service management at a large scale. The later address the human factor in building interactive

systems but do not focus on user collaboration. Our project functions as a bridge between these two approaches. We aim to support trustworthy collaboration and we address collaborative data and service management at a large scale where the user plays a central role. Our project considers the business layer, its relation to the IT layer and to the users. Regarding the tools that we use to conduct our research, we are commonly looking towards providers of formal models that can be leveraged to validate the models that we are designing. For instance, we have a long and continuing history of collaboration with the Cassis project.

PAST AND CURRENT COLLABORATIONS INSIDE INRIA

REGAL http://www.inria.fr/equipes/regal

The domain of research of REGAL is large-scale distributed computing systems with a focus on theoretical aspects related to system level. We have a complementary approach where we design consistency maintenance algorithms and collaborative systems from a user'S point of view and we evaluate our systems with real collaboration traces and with users. REGAL is our regular partner in ARC RECALL (2006–2007), ANR ConcoRDanT (2010–2014) and ANR STREAMS (2010–2014).

ASAP http://www.inria.fr/equipes/asap

ASAP team focuses on the design of large-scale distributed systems from two major directions: models and abstractions for distributed computing and user-centric fully decentralized systems and applications. Although we share the same vision of a decentralized approach that does not rely on a central authority, we have a complementary research direction focusing on data management from a collaborative point of view and on a service composition approach. ASAP is our partner in ANR STREAMS (2010–2014).

ZENITH http://www.inria.fr/equipes/zenith

Zenith team focuses on large scale scientific data management addressing issues of data integration, scientific workflows, recommendation, query processing and data analysis in clusters, peer-to-peer and cloud environments. We have a complementary approach focusing on collaborative data management and service composition. Zenith was our partner in ANR XWiki Concerto (2006–2009).

SPIRALS http://www.inria.fr/equipes/spirals

SPIRALS (formerly ADAM) deals with the concepts and tools for the adaptation of applications and middleware in distributed multi- scale environments. They rely on software engineering techniques, such as Component-Based and Aspect-Oriented Software Development and Context-Aware Computing, while we rely on a service-oriented approach focusing on trustworthy data-centric service composition. We were involved in the Galaxy ADT whose aim was to provide an open SOA Agile Platform.

ALGORILLE http://www.inria.fr/equipes/algorille

We are setting up collaboration with the Algorille project-team team on the use of process engine for large scale distributed systems experimentation. This is ongoing and it is related to our investigation on the question of large scale execution of process.

MADYNES http://www.inria.fr/equipes/madynes

As part of the ICT Labs initiative, we have contributed to a common project submission Crowdsourcing processes for Smart cities. They have competencies on mobile operation and deployment that we currently don't have and that are useful for our project.

CASSIS http://www.inria.fr/equipes/cassis

The aim of the project is the design and the development of tools for checking the safety of systems with an infinite number of states. Their analysis of systems is based on a symbolic representation of the sets of states as formal languages or logical formulas. Safety is obtained by automated proofs, symbolic exploration of models, or tests generation. We have a long history of collaboration with the CASSIS project-team. This is reflected by numerous publications both on safe service composition and replications algorithms. We are also partners in the ANR STREAMS project.

ORPAILLEUR http://www.inria.fr/equipes/orpailleur

We are collaborating with the Orpailleur project-team in the context of the ANR KolFlow project (2010–2014). We are providing our expertise for the collaborative construction of knowledge base or ontology.

8.3 INRIA PROJECT WITH INTERSECTING RESEARCH PROBLEMS

KERDATA http://www.irisa.fr/kerdata KerData is looking at distributed data management for data-oriented applications that handles large amount of unstructured data. They are concerned by the infrastructure layer whereas we are working at the Software as a Service Layer (SaaS) and platform. Their requirements are different than ours since we are considering the users in the loop.

ARLES https://www.rocq.inria.fr/arles/ The Arles team conducts research that is similar to ours to some extent. They have a focus on the formalizing and engineering of software architectures for distributed systems and service architecture. They are also considering the problem of scale with service choreography. Still, they have are more focused on the middleware layer than on the business one. There view on service composition is also different than the one we advocate.

8.4 NATIONAL AND INTERNATIONAL RELATIONSHIPS

The originality of the work is to leverage competencies in the domain of large scale optimistic data sharing and in the domain of service composition with a strong emphasis on the social dimension and on trust. Our goal is to try to keep our work vertically aligned across theoretical, social and technological dimensions.

In the international research landscape (our social network), our ambition is to become a result oriented hub in the domains of data sharing and of service composition and as a potential source of challenging test/ scenarios for theoretical and experimental domains. To our knowledge, no other project or team combines at the same time our skills in CSCW, data sharing and service

composition. Still, we maintain strong relationships with other teams in Europe and in the World on the domain of Service Computing : Boualem Bennatallah from University of New South Wales in Australia, Wim Van der Aalst from Technical University of Eindhoven, Sharam Dustdar from University of Vienne, Fabio Casati from University of Trento. In the domain of Collaborative editing we are organising an yearly workshop with the main contributors in the domain such as David Sun from University of California, Berkeley, Chengzheng Sun from Nanyang Technological University and Ning Gu from Fudan University. We also collaborate with the Nuño Pregui ca from Universidade Nova de Lisboa, Valerie Shalin from Wright State University and Weihai Yu from University of Tromsø..

8.5 BUSINESS AND TRANSFER OPPORTUNITIES

Bonita, the flagship product of Bonitasoft, began 10 years ago with Ecoo project-team, Score and Coast ancestors. Now that the product is a recognized success in the BPM market, there is an urgent need to consider its future. The work that we have done and that we plan to do on coordination, on trust and on service management will touch future versions of Bonita in many ways. Conversely, our work will benefit from a direct cooperation that will nurture us with new use cases and current trends in process management and in social coordination.

The availability of Cloud infrastructures is a transfer accelerator for the kind of services that we can derive- from our research results. The use of cloud services generated more input for our work. Following this strategy, we will publish and maintain services rather than publish code. It will require less man power to deploy and maintain than the delivery of software libraries or systems that we have done until now.

For years, the Ecoo and Score teams have developed software and frameworks (Bonita, LibreSource, QualiPSo) that were difficult to maintain and deploy efforts to simplify these tasks. With a service oriented perspective, in a Cloud based IT world, it becomes much easier to deliver and maintain services than software. We do not need a sophisticated platform to make our work available to the world. Since our goal is to design and validate composable services, aimed at service provider, for data management, trust management and service management in general, we will be able either to provide them as standalone service delivered on a public Platform as a Service (PaaS) or we will be able with to deploy them on dedicated service delivery platform. This is also in line with our support of the Open Web and Open Data initiatives.

We expect to get patentable results and/or services that companies could incorporate in their existing solutions. We are confident in growing service market comparable to the current dynamic market for mobile applications. However, as usual in the context of services on the web, what matters is to provide value. If this value is recognized and well marketed, the business follows. Again, the Infrastructure as a Service (IaaS) allows deploying services with very little start-up capital. Our expectation is that if we are able to provide a service that would help to raise confidence in data management services and their composition, we could transfer this service with new business opportunities.

With sufficient resources, a long term outcome is a platform for the composition and the delivery of services whose trustworthiness could be assessed depending on stakeholder defined criteria, based our model. Thus, we are placing ourselves in the area of services for services management.

An efficient and open cloud based and large scale collaborative environment could also be an outcome of the project as a proof of concept of some of our services. The market for

these environments is still nascent and the demand is high but people are still searching for appropriate and proven solutions. The importance that we will give to trust should be a competitive advantage in this area.

8.6 SOCIETAL OPPORTUNITIES

Our results and services will support deployable solutions for collaboration among communities, outside of the control of centralized privately owned foreign companies or of state agencies. This will depend of course on the result that we will obtain in pair-to-pair environments. The "Arab spring" is a vibrant example of a situation where the availability of a lightweight solution to support large scale trustworthy collaboration – can be set up out of the control of government and of large corporations.–

Other applications that we are also considering today are crowdsourcing applications that require the execution of new kind of process (social or crowdsourcing process) combining traditional business process patterns and more scientific patterns. We are considering crowdsourcing scenarios in the context of Smart Sustainable Cities or in the context of Crisis Management. In both of these contexts, leveraging the workforce of dedicated communities with the right services can help finding solutions to problem that would be out of reach otherwise to city stakeholders.

9 MEMBERS SHORT CVS

KHALID BENALI (ASSOCIATE PROFESSOR, HDR, UNIVERSITÉ DE LORRAINE) has an expertise in distributed and cooperative systems, and integration/interoperability of virtual enterprises. He received his PhD in Computer Science from Nancy 1 University, France, in 1989 and his HDR (Habilitation à Diriger les Recherches) from Nancy 2 University, France, in 2004. His lectures are mainly in the area of distributed systems and object oriented analysis and design for postgraduates and he is responsible of Master 2 MIAGE SID in France as well as in a shared diploma with IGA in Morocco. He participated actively to several french and european project (ESPRIT, IST, Network of excellence), published in international conferences like CoopIS, BPM and has been invited editor for several french journals like (Ingénierie des systèmes d'information-RST) or PC member and reviewer of conferences like I-ESA, IWEI or I3E.

GÉRÔME CANALS (ASSOCIATE PROFESSOR, UNIVERSITÉ DE LORRAINE) has an expertise in distributed collaborative systems with a particular focus on awareness mechanisms, collaborative construction of knowledge and mobile collaborative systems. He first worked on advanced transaction models and systems for the coordination of long termed and collaborative design activities. He then was focused on group awareness and knowledge sharing. He supervised a phd thesis that proposed a context-based awareness system for mobile collaboration, and a phd thesis that proposed a sharing model for an ontology editor. The work resulted in the proposition of a divergence awareness mechanism and a collaborative version of an ontology editor (Co-protégé). He then applied this work on awareness to the context of P2P wiki and proposed a concurrency awareness system for a decentralized wiki. He is actually

working on the building of a distributed collaborative space of the collaborative construction of knowledge, with a particular focus on test approaches for the validation of knowledge edition.

Gérôme is a co-founder of the UbiMob french conference. He participated in several collaborative research projects in past years, and particularly the ANR XWiki Concerto in which he was the principal investigator, the DGCIS Wiki 3.0, and the ongoing ANR KolFlow. Gérôme is actually head of the computer science departement at the Nancy-Charlemagne institute of technology (IUT).

FRANÇOIS CHAROY (FULL PROFESSOR, UNIVERSITÉ DE LORRAINE) has an expertise in Business Process Modelling, Service Oriented Architecture, Information System Security, workflow management. He created the first version of Bonita, a cooperative business process management system now supported by BonitaSoft. He also collaborate with SAP from 2006 to 2012 on security in workflows and on the application of process management in extreme condition like crisis management. More recently, he conducted some research on crowdsourcing for Smart Cities, as a way to experiment process management at new scales. He has published several articles in international journals (Enterprise IS Journal, IJ of Software Engineering and Knowledge Engineering, AI & Society), and conferences (WISE, CAiSE, BPM, ICWS, SCC, CoopIS, CollaborateCom, ...), he was involved in several international and European projects. He has been Co-Chair of the CAISE PhD symposium 2011, and Program Committee member of many conferences (BPM, CoopIS). He is member of the editorial board of Service Oriented Computing and Applications. He reviewed in 2013 papers for IEEE Transactions on Services and Network Management, IEEE Transactions on the Web, Web Services Handbook, Journal of Intelligent Information Systems, and Software and Systems Modeling journals.

CLAUDE GODART (FULL PROFESSOR, UNIVERSITÉ DE LORRAINE) has lead and conduct researches about business process management and Web services composition in complex settings: flexible process models, transactional processes, Web services composition and governance with QoS properties, trusted execution of processes in clouds. He is author and co-author of more than 200 papers, adviser of around 30 PhD theses, and participated in many contractual activities (private – including the Hitachi/INRIA chair in 1999 –, ANR, european – NoE, IP, ...–, regional). He is member of the editorial board of several journals (IEEE Transactions on Service Computing, ...) and of conference program committees (IEEE ICWS, IEEE Cloud, BPM, ...).

CLAUDIA IGNAT (INRIA RESEARCHER, CR1) has an expertise in distributed collaborative systems with a focus on consistency maintenance, group awareness, security and trust issues and more recently user studies. She developed several algorithms for maintaining consistency based on operational transformation and CRDT approaches for textual, hierarchical and graphical documents and wiki pages. She proposed an awareness mechanism in the context of software engineering, collaborative editing of textual documents and of web pages. She co-supervised a PhD thesis that proposed a contract-based collaboration model as a support for a "soft" security mechanism in collaborative editing where access is given first to data without control but with restrictions that are verified a posteriori. In the context of this work a mechanism was proposed for establishment of trusted logs relying on hash-chain based authenticators. Recently, she started working on the evaluation with users of collaborative approaches and systems, particularly on the evaluation of real-time constraints in the collaboration.

Claudia is an associate editor of Journal of Computer Supported Cooperative Work, a regular

PC member of the GROUP, CDVE and ICEBE conferences and organizer since 2003 of the International Workshop on Collaborative Editing. She is the principal investigator of the Inria associated team USCoast and she participated to the DGCIS Wiki 3.0 project (responsible of the Inria partner) and ANR STREAMS and ConcoRDanT.

GÉRALD OSTER (ASSOCIATE PROFESSOR, UNIVERSITÉ DE LORRAINE) has an expertise in distributed collaborative systems with a focus on content replication mechanisms and their applicability. During his PhD, he worked on verification of correctness of a family of optimistic replication mechanisms (operational transformation) dedicated to collaborative editing. He proposed a framework based on a automated theorem prover and several sets of verified transformation functions for multiple data types. He worked on the design and the implementation of a universal file synchronizer. He is one of the father of the CRDT approach as he participated in the design of the WOOT algorithm which initiated researches on these distinctive data structures. He is currently investigating the limitations and the applicability in diverse domains of these novel replicated data structures.

Gérald is the principal investigator of the ANR STREAMS project and currently involved in the ANR ConcoRDanT project and USCoast associate-team. He participated in several technology transfer oriented projects such as RNTL LibreSource, FP6 IP QualiPSo, ANR XWiki Concerto, DGCIS Wiki 3.0.

OLIVIER PERRIN (FULL PROFESSOR, UNIVERSITÉ DE LORRAINE) has an expertise in Business Process Modeling for virtual organizations; Web services compositions, workflow management, and enterprise application & data integration, covering the process life cycle from modeling to execution including compliance purposes. He proposed the DISC event-oriented framework which was a unified declarative framework designed to reconcile the process design, verification and monitoring of services compositions. DISC provides a flexible and highly expressive composition design that can accommodate various aspects such as data relationships and constraints; services dynamic binding, compliance regulations, security or temporal requirements. Further, the DISC framework allows for instantiating and verifying the composition design and for monitoring the process while in execution. He is currently working on enhancing security aspects (authentication, authorization and trust management) within a open Platform as a Service in order to allow collaborative activities and services to be handled between various partners. Olivier Perrin has published several articles in international journals (Data & Knowledge Engineering, Distributed and Parallel Databases, Journal of Data Management, ...), and conferences (World Wide Web, ICSOC, WISE, CAiSE, BPM, ICWS, SCC, CoopIS, ...), international workshops on Governance, Risk and Compliance - Applications in Information Systems and he was involved in many international and European projects. He has been recently Co-Chair of the ICSOC PhD symposium 2012, and Program Committee member of ICSOC 2012/2013, WCE (Workshop on Capability based Engineering) workshop at WISE 2012, GRCIS workshop 2012. He was reviewer for BPM 2012. He reviewed in 2013 papers for IEEE Transactions on Services Computing, IEEE Transactions on Software Engineering, IEEE Transactions on the Web, Web Services Handbook, Journal of Intelligent Information Systems, and Software and Systems Modeling journals.

PASCAL URSO, ASSOCIATE PROFESSOR, UNIVERSITÉ DE LORRAINE). His research interests include distributed systems, data replication, collaborative systems, P2P computing and automated theorem proving. He has a Master's (1998) and a PhD (2002)

degree from Université de Nice – Sophia Antipolis. His PhD dissertation was about automated inductive theorem proving and its applications. He worked as a post-doctoral fellow (2003) at the University of Namur (FUNDP). Pascal Urso participated to several projects including ANR projects, NoE-Interop and EU FP6 Qualipso, reviewed international journal papers and organized a French-speaking conference (UbiMob 2013). He has supervised 1 PhD thesis and 3 MSc theses. He pioneered algorithms for replicating structured mutable data in large-scale distributed systems. He had conducted several works about implementing and evaluating optimistic replication algorithms in the context of large-scale and real data. He is now responsible in the SCORE group for ANR Concordant and Syncfree STREP projects. He is co-supervising 2 PhD thesis (M. Ahmed-Nacer and J. Martori i Adrian, just starting). His current researches focus on designing, evaluating and establish correctness of eventually consistent architectures and algorithms. When evaluating solutions, performance but also user satisfaction is measured. Pascal Urso is also responsible of the "Security, Services, Systems and Network" track of MSc in Computer Science at the University de Lorraine.

SAMIR YOUCEF, ASSOCIATE PROFESSOR, UNIVERSITÉ DE LORRAINE). He received PhD degrees in computer science in 2009 from University Paris-Dauphine. His main research interest is about performance evaluation of Web services and matching and scheduling processes in Cloud computing context. During his PhD, he worked on developing methods and tools for composite Web services performance evaluation. To achieve this goal, our approach is driven by three aspects, namely, the response time exact computation, the response time bounds and the taking account of Web services discovery and selection according to their QoS. For the first aspect, he proposed analytical formulas for the exact computation and the analysis of the various BPEL standard constructors response times. For the second aspect, upper bounds for composite Web services response times have been provided. The analysis is based on continuous time Markov chain (CTMC) and coupling processes. For the third aspect, an extension of the conventional Web services architecture in order to take into account the QoS in their discovery and selection has been proposed.