

**DOSSIER DE CANDIDATURE AUX FONCTIONS DE MAÎTRE DE  
CONFÉRENCES**

Poste N°1452, référence galaxie 4076, CNU 27  
Equipe d'intégration : GDD, LINA, Université de Nantes

**M. STÉPHANE MARTIN**

ATER

à

l'Université de Lorraine (UL)  
au laboratoire lorrain de recherche en informatique et ses applications  
(LORIA)

Docteur en Informatique de l'Université de Provence  
Laboratoire d'Informatique Fondamentale (LIF)





*Ce document doit être joint au dossier transmis à  
l'intention du comité de sélection*

DÉCLARATION DE CANDIDATURE AU RECRUTEMENT SUR UN EMPLOI DE MAITRE DE CONFERENCES  
(Campagne 2012)

(décret n° 84-431 du 6 juin 1984 modifié)  
Authentification : f6b9f993e5f8b807af83e4c738eed5f8 (1332752003878)

adressée au chef d'établissement de : UNIVERSITE DE NANTES

Poste(s) : n°4076

Publié le : 23 février 2012 à 00:02

Section(s) C.N.U : 27 (Informatique)

Profil : Informatique

Localisation : NANTES

Article 26-I-1

Chaire : Non

Je soussigné(e) M.

**Nom patronymique : MARTIN**

Nom d'usage ou marital :

Prénom : STEPHANE

Date et lieu de naissance : 07/02/1984 - SAINT-DENIS

Nationalité : Française

N° de qualification : 12227226076

Adresse postale et électronique à laquelle seront acheminées toutes les correspondances		
14 RUE DE LA BERGAMOTE		
Code postal : 54100	Ville : NANCY	Pays : FRANCE
Téléphone : 0611438009	Télécopie : 0383413079	
Adresse électronique : but2_ene@yahoo.fr		

Fonctions et établissement actuel : ATER n'ayant pas la qualité de fonctionnaire-UNIVERSITE HENRI POINCARÉ

**Date de création :**

23/02/2012 à 18:02

**Date de dernière modification :**

26/03/2012 à 09:03

**Titres universitaires français :**

Doctorat

**Diplôme le plus récent : Thèse**

Titre : Edition collaborative des documents semi-structurés

Date de soutenance : 08/09/2011

Lieu de la soutenance : UNIVERSITE DE PROVENCE

Mention : L'université de provence n'en délivre plus.

Directeur : DENIS LUGIEZ

Composition du jury : VERONIQUE BENZAKEN

PASCAL MOLLI

MARC SHAPIRO

MICHAEL RUSINOWITCH

REMI MORIN

DENIS LUGIEZ

**Activités en matière d'enseignement :**

ATER(2010-2012)

CM/TD/TP, L3 info reseaux

CM/TD/TP, Dut SRC architecture des ordi.

TD/TP, DUT info, initiation en java

TP, Esial prog. d'application rép.

TP, Web.

CM/TD/TP,L1 info, Initiation informatique

TP, L1 Bio, Initiation informatique

TD/TP,L3, Architecture et systemes d'exploitation

Moniteur(2007-2010)

TP,L3 info, Stage C

TD/TP,M1 info, Sys. paralleles et distribues

**thème de recherche et mots clés :**

Collaborative Editing, Peer-to-Peer, Semistructured document, Commutative Replicated Data Type (CRDT), Operational Transformation, Types for XML, Distributed Systems, Software Engineering.

**Activités en matière d'administration et autres responsabilités collectives :**

Représentant des doctorants à l'école doctorale,

Vice-Président du cyberfoyer : Association étudiante 1901, proposant du café, des boissons, des viennoiseries, ... et des ordinateurs en libre accès (maintenance, achat, négociation boulangerie,...).

**Autres titres et diplômes :**

**Travaux, ouvrages, articles, réalisations :**

2012: Stéphane Martin, Mehdi Ahmed-Nacer , Pascal Urso, Abstract unordered and ordered trees CRDT, Delivrable ANR + art. soumis PODC'12

2010: Collaborative edition on typed document - CDVE'10

Scalable XML Collaborative Editing with Undo - CooPIS'10 Stéphane Martin, Pascal Urso et Stephane Weiss

2009 : Avoiding Conflicts is Better than Solving Conflicts - AC'09

**Liste des pièces à fournir par le candidat :**

pièces obligatoires

Il est impératif de consulter le site internet de l'Université de Nantes afin de s'assurer des pièces à fournir. Le dossier de candidature est à envoyer prioritairement sous format électronique, sur un seul fichier de format PDF.



déclare faire acte de candidature sur l'emploi ci-dessus désigné :

Fait à Nancy

le 26/03/2012

Signature

A handwritten signature in black ink, appearing to be 'L. M. G.', written over a horizontal line.



# Attestation de réussite du Doctorat

---



## ATTESTATION DE RÉUSSITE AU DIPLÔME

Le Président de l'Université de Provence certifie que  
**Le Doctorat d'Informatique**  
a été décerné, conformément à l'arrêté du 7 août 2006, à  
**Monsieur STEPHANE MARTIN**  
né le 07 Février 1984 à SAINT-DENIS  
au titre de l'année universitaire 2010-2011 avec la mention Très honorable

Titre des travaux : Edition Collaborative des documents semi-structurés  
Date de soutenance : 08 Septembre 2011  
Etablissement de soutenance : UNIVERSITE AIX-MARSEILLE 1  
Ecole doctorale : Mathématiques et informatique de Marseille (184)  
Directeur : M. DENIS LUGIEZ

### Membres du jury

Monsieur RUSINOWITCH MICHAEL - Membre du jury - Directeur de recherche - INRIA  
Monsieur MOLLI PASCAL - Rapporteur du jury - HDR - UNIVERSITE NANTES  
Madame BENZAKEN VERONIQUE - Rapporteur du jury - HDR - UNIVERSITE PARIS 11  
Monsieur LUGIEZ DENIS - Membre du jury - Professeur ou assimilé - UNIVERSITE AIX-MARSEILLE 1  
Monsieur MÔRIN RÉMI - Président du jury - Professeur ou assimilé - UNIVERSITE AIX-MARSEILLE 2  
Monsieur SHAPIRO MARC - Membre du jury - Directeur de recherche - UNIVERSITE PARIS 6

Fait à MARSEILLE, le 08 Septembre 2011

Pour le président et par délégation  
Le Vice-Président du Conseil Scientifique





# Table des matières

<b>1 Curriculum vitæ</b>	<b>1</b>
<b>2 Recherche</b>	<b>3</b>
2.1 Master Recherche . . . . .	3
2.2 Thèse . . . . .	4
2.3 Recherches actuelles . . . . .	5
2.4 Projet d'intégration . . . . .	6
2.5 Publications . . . . .	7
<b>3 Enseignement</b>	<b>9</b>
3.1 Synthèse des enseignements effectués . . . . .	9
3.2 Expériences et visions . . . . .	9
3.3 Projet d'enseignement et intégration . . . . .	10
3.4 Détail des enseignements . . . . .	12
<b>4 Charges Collectives</b>	<b>17</b>
<b>5 Recommandations professionnelles</b>	<b>19</b>
<b>A Documents attachés</b>	<b>21</b>
A.1 Copies des lettres de recommandations . . . . .	22
A.2 Procès verbal de soutenance . . . . .	30
A.3 Attestation d'enseignements . . . . .	35
A.4 Attestation directeur de l'école doctorale (Charges collectives) . . . . .	37
A.5 Rapports de présoutenance de Véronique Benzaken . . . . .	39
A.6 Rapports de présoutenance de Pascal Molli . . . . .	43



# Stéphane Martin

*Attaché d'Enseignement et de  
Recherche - Docteur*



Curriculum vitæ

---

## INFORMATIONS PERSONNELLES

Date de nais. 07 février, 1984  
Lieu de nais. St-Denis (93), France  
Nationalité Française  
Pacsé  
Adresse 14 rue de la bergamote 54100 Nancy.  
Téléphone Pro. +33 3 54 95 86 29  
Téléphone Pers. +33 3 54 12 43 17  
Mobile +33 6 11 43 80 09  
Email stephane.martin@loria.fr  
Site Web <http://www.loria.fr/~stemarti>

---

## FONCTIONS RÉCENTES

2011-2012 Attaché temporaire d'enseignement et de recherche auprès de l'UFR Science et de l'IUT de st-Dié, à l'Université Henri Poincaré  
2010-2011 Attaché temporaire d'enseignement et de recherche auprès de l'UFR Mathématiques Informatique et Mécanique, à l'Université de Provence  
2007-2010 Allocataire de recherche MENRT et Moniteur auprès de l'UFR Mathématiques Informatique et Mécanique, à l'Université de Provence

---

## DIPLÔMES ET FORMATIONS

2007-2011 **Doctorat en Informatique**, de l'Université de Provence Marseille, préparé au Laboratoire d'Informatique Fondamentale (LIF).  
Titre «Édition collaborative des documents arborescents»  
Superviseur M. Denis LUGIEZ, Professeur à l'Université de Provence  
M.<sup>me</sup> Véronique BENZAKEN, Professeur à l'Université Paris-Sud XI (Rapporteur)  
M. Pascal MOLLI, Professeur à l'Université de Nantes (Rapporteur)  
Jury M. Marc SHAPIRO, Directeur de Recherche INRIA  
M. Michaël RUSINOWITCH, Directeur de Recherche INRIA  
M. Rémi MORIN, Professeur à l'Université de la Méditerranée  
2007 **Master 2 recherche en Informatique**, à l'Université de Provence – *Mention Bien*  
Titre Mémoire «Vers la vérification automatique des services Web».  
Superviseur M. Denis LUGIEZ, Professeur à l'Université de Provence

- 2006 **Master Informatique**, à l'Université de Provence – *Mention Assez Bien*
- 2005 **Licence Informatique**, à l'Université de Provence – *Mention Très Bien*
- 2004 **Deug Mathématiques et Informatique Appliqués aux Sciences**, à l'Université de Provence
- 2002 **Baccalauréat Section S Option Sciences de l'Ingénieur**, au lycée des Remparts à Marseille





## 2.1 Master Recherche

---

Mon stage de master portait sur «La vérification automatique des services Web». Les services Web sont des programmes qui permettent de réaliser des transactions via internet, comme par exemple : réserver des titres de transport, commander des produits, ou faire des consultations en ligne, etc. Leur succès s'est construit en s'appuyant sur un ensemble de standards tels que le langage de description «Web Services Description Language» (WSDL), de transmission «Simple Object Access Protocol» (SOAP) ou de composition «Business Process Execution Language» (BPEL). Malgré de nombreux avantages avérés des services web, ils restent confrontés à de nombreux problèmes tels que des problèmes de sécurité. Il est donc nécessaire de développer des méthodes et des outils afin de garantir un niveau de sécurité suffisant à leur utilisation. Ce problème recouvre des notions très diverses depuis la non-divulgation de données secrètes (numéro de carte bancaire par exemple), à l'authentification en passant par la qualité de service etc. La difficulté du problème vient d'une part de la modélisation qui est loin d'être facile et d'autre part de l'élaboration de méthodes de vérification automatique. Ce domaine fait l'objet de nombreuses recherches tant au niveau académique que privé (Microsoft, SAP,...).

**Problématique :** Comment vérifie-t-on la préservation du secret pour un modèle de services basés sur le protocole SOAP ? Dans ce modèle, les services sont des protocoles faisant intervenir des agents. Ces agents communiquent par la réception/envoi de messages au format SOAP. Ce modèle est proche de ce qui existe pour les protocoles cryptographiques de Dolev Yao<sup>1</sup>. Cependant, la modification essentielle est que les messages sont des données semi-structurées respectant la norme SOAP : un message est une enveloppe contenant un body (obligatoire) et un header (facultatif) qui contiendra en outre les aspects de sécurisation du message. Un protocole met en oeuvre une politique de sécurité pour s'assurer que le message est conforme. Nos travaux sont basés sur la WS-SecurityPolicy normalisée par le W3C.

**Contributions :** Après un état de l'art sur les services et les normes, nous avons abordé le problème sous trois aspects :

1. Nous avons ajouté des contraintes de langages réguliers dans la vérification des protocoles cryptographiques (les messages sont des termes et pas des données semi-structurées). Ce qui permet de définir les protocoles cryptographiques avec des contraintes de régularité. Les langages réguliers sont une première approche pour l'étude de politiques de sécurité.
2. Nous avons obtenu un résultat de localité du système de déduction qui exprime le pouvoir de l'attaquant, dans le cadre d'échange de données semi-structurées et non des termes. Cette propriété établit que tout ce qui est nécessaire pour obtenir un message à l'aide des règles (DY), est «écrit» dans la connaissance initiale et le message. Elle montre la décidabilité du problème du secret.

---

<sup>1</sup>D. Dolev & A.C. Yao. On the security of public key protocols. In Proceeding of the 22nd Symp. on Foundations of Computer Science, Pages 350-357, Nashville, Tennessee, USA, 1981. IEEE Computer Society Press.

3. Nous avons élaboré et prouvé un système de déduction permettant de résoudre le problème de la sécurité des protocoles XML. Chaque résultat est une extension de ce qui a été fait dans le cas classique et a demandé une extension de l'algorithme utilisé. Nous nous sommes basés sur les règles de Comon-Lundh<sup>2</sup>.

Pour plus amples information le mémoire est disponible sur ma page web : <http://www.loria.fr/~stemarti/recherches>

## 2.2 Thèse

---

Le titre de ma thèse est «Edition collaborative des documents semi-structurés», et elle a été dirigée par Denis Lugiez, Professeur à l'université de Provence au Laboratoire d'Informatique Fondamentale.

Les éditeurs collaboratifs permettent à des utilisateurs éloignés de collaborer à une tâche commune, par exemple : l'utilisation d'un agenda partagé ou la réalisation de logiciels. Ce concept est né avec NLS<sup>3</sup> en 1968 de Egelbart, suivit de SCCS<sup>4</sup> en 1972. L'utilisation de ces systèmes collaboratifs est devenue courante, que ce soit pour le développement d'un logiciel, l'écriture d'une encyclopédie ou d'un document. (ex: Git, Wikipedia, SVN, google doc, ...). Les réseaux pair-à-pair connaissent un intérêt grandissant grâce à leur facilité de déploiement, leur flexibilité, leur faible coût, etc. D'un autre côté, le format XML est devenu une référence pour la manipulation et l'échange de documents. Notre travail vise à la réalisation d'un éditeur collaboratif P2P pour l'édition de documents semi-structurés qui sont une abstraction du format XML et d'autres format tel que Json<sup>5</sup>. Le problème est difficile et de nombreuses propositions se sont révélées erronées ou ne passant pas à l'échelle.

**Problématique :** La problématique de cette thèse est d'assurer la convergence de données distribuées. Pour cela nous avons exploré deux modèles basés sur des opérations : le modèle des transformées opérationnelles<sup>6</sup> et le modèle CRDT (Commutative Replicated Data Type<sup>7</sup>) sont appliqué sur différentes structures de données arborescentes. L'objectif est de supporter les opérations de base (ajout, suppression et ré-étiquetage) tout en garantissant la convergence du processus d'édition.

J'ai commencé cette thèse en 2007 à partir des travaux d'Abdessamad Imine (LORIA à Nancy) sur les transformées opérationnelles pour du texte<sup>8</sup>. Les exemples d'impossibilités issue de cette thèse m'ont fait comprendre qu'il fallait changer la méthode d'ordonnement des éléments du document.

### Contributions :

1. J'ai commencé par écrire des transformées opérationnelles ayant les propriétés de convergence pour les arbres non ordonnés. Par la même occasion, j'ai trouvé et démontré quelques cas d'impossibilités, ce qui est innovant pour le domaine[6].
2. J'ai proposé un algorithme générique pour l'approche CRDT basé sur une notion d'indépendance dans la structure de données. L'intérêt de l'approche est d'avoir une propriété de convergence plus simple que celle dans le modèle des transformées opérationnelles : «Si deux opération sont indépendantes alors elles peuvent commuter». J'ai proposé une structure de données d'arbre ordonné avec les opérations et la relation de dépendance permettant la convergence de cet algorithme [1].

---

<sup>2</sup>Hubert Comon-Lundh, Résolution de contraintes et recherche d'attaque pour un nombre borné de sessions, 5 février 2004.

<sup>3</sup>Engelbart, D., Study for the development of Human Augmentation Techniques. Final Report, July 1968. Sections 4 and 5.

<sup>4</sup>Source Code Control System (SCCS), est un système de gestion de versions.

<sup>5</sup>JavaScript Object Notation

<sup>6</sup>Chaque opération est transformée vis à vis des concurrentes.

<sup>7</sup>les opérations concurrentes sont commutatives entre elles.

<sup>8</sup>Conception Formelle d'Algorithmes de Réplication Optimiste. Vers l'Édition Collaborative dans les Réseaux Pair-à-Pair. Université Henri Poincaré, Nancy, décembre 2006.

Avec la collaboration de Pascal Urso et de Stéphane Weiss, nous avons adapté ces résultats avec ceux de logoot undo<sup>9</sup> pour les étendre aux fichiers XML [3].

3. J'ai étendu ces travaux afin de réaliser l'opération de déplacement d'un sous-arbre. Pour cela, j'ai modélisé les graphes fonctionnels dans le modèle des transformées opérationnelles et le modèle commutatif. Le problème de cette modélisation est la création de cycle. Dans un graphe fonctionnel, un cycle n'est plus connexe à la racine. Il en résulte une perte d'informations. Ce problème reste une question ouverte que nous tentons de résoudre actuellement. J'ai écrit par la même occasion une version généralisée pour les multigraphes. Ces multigraphes sont utilisés par exemple pour faire des cartes routières.

Ces derniers résultats ne sont pas encore publiés.

4. J'ai développé un prototype en java permettant de montrer que mon algorithme était utilisable, et que le jeu d'opérations proposé répond à certains problèmes de l'édition collaborative comme le déplacement d'un paragraphe pendant sa modification.
5. Enfin, j'ai en partie, résolu les problèmes de typage dans les données distribuées. J'ai utilisé une notion de vue équivalente à celle présente dans les bases de données. J'ai réutilisé les derniers résultats sur la convergence d'arbres répliqués que l'on répare à l'aide d'un algorithme de réparation d'arbre<sup>10</sup>. Cet algorithme nous donne une vue utilisateur respectant le typage. Toutes les opérations générées par l'utilisateur sur cette vue sont adaptées pour être compatibles avec la vue globale. Le fait intéressant est que le système garde sa propriété de convergence. Ce résultat a été publié à CDVE[2] en 2010. Bien que fondamentaux dans l'édition collaborative, le déplacement d'arêtes et le typage d'arbres ont très peu été abordés dans la littérature.

## 2.3 Recherches actuelles

Cette année, j'ai intégré l'équipe SCORE (Services and Cooperation) du LORIA à Nancy. J'essaye de résoudre certains problèmes restés ouverts sur les objets répliqués.

Dès mon arrivée, j'ai collaboré avec Pascal Urso et Medhi Ahmed-Nacer qui réfléchissaient sur les différentes façons de résoudre des conflits sur des arbres dans le modèle commutatif.

Nous avons revisité les ensembles CRDT pour décrire des arbres. Un arbre peut être vu soit comme un graphe avec un ensemble d'arêtes et un ensemble de nœuds ou un ensemble d'arêtes où on en déduit l'ensemble de nœuds; soit comme un ensemble de chemins. Lors d'exécution d'opérations concurrentes, il est possible, dans les deux premières modélisations, que des cycles apparaissent. Dans ces trois modélisations, il est possible qu'une partie de l'arbre disparaisse en devenant non connexe à la racine. Nous avons exhibé plusieurs façons de générer une vue d'arbre à partir de ces ensembles. Nous avons aussi montré qu'il est possible de composer ces éléments CRDT de base et les vues pour combiner d'autres CRDT. Cette collaboration a donné lieu à un livrable pour l'ANR concordant [4] <http://concordant.lip6.fr/> et à un article que nous avons soumis à PODC'12.

Nous pensons pouvoir résoudre les cycles dus à l'opération de déplacement de la même manière. Nous avons commencé la rédaction d'un article pour un journal concernant le déplacement et la résolution de cycles.

Dans un prochain article, nous étudierons le cas d'un typage à l'aide d'une DTD. L'intérêt d'utiliser une DTD est d'avoir une propriété de localité sur le type pour définir une réparation locale à chaque modification de façon incrémentale. Les types XML Schema et relax NG peuvent avoir des conditions sur des nœuds à d'autres endroits de l'arbre (un nœud peut dépendre d'un autre sur une autre branche) ce qui rend les réparations difficiles. La méthode que nous mettons au point sera beaucoup plus efficace que celle décrite dans ma thèse et publiée à CDVE [2]. Dans un second temps, nous explorerons une façon de limiter l'expressivité de XML Schema ou relax NG afin de limiter le coût des réparations.

<sup>9</sup>Stéphane Weiss, Pascal Urso, and Pascal Molli. Logoot-undo : Distributed collaborative editing system on p2p networks. IEEE Transactions on Parallel and Distributed Systems, 99(Pre-Prints), 2009.

<sup>10</sup>S. Staworko and J. Chomicki. Validity-sensitive querying of XML databases. In EDBT Workshops (dataX), pages 164–177. Springer LNCS 4254, 2006.

Je participe à un projet STREP sur la maintenance de la cohérence d'objets extrêmement répliqués. Actuellement tous les systèmes pair-à-pair étant à l'échelle d'internet (torrent, magnet, Kad,...) partagent des objets non mutables. Le but de ce projet est d'apporter des formalismes, des méthodes et modèles de programmation pour partager des objets modifiables avec différents comportements. Le système devra passer à l'échelle d'internet.

Tout ceci s'intègre dans mon projet de recherche à savoir: comment faire des systèmes massivement collaboratifs passants à l'échelle en étant pair-à-pair? C'est à dire sans grosse structure couteuse tel qu'un cloud.

## 2.4 Projet d'intégration

---

Un système collaboratif possède plusieurs aspects. Dans ma thèse et dans l'équipe Score, j'ai surtout regardé la cohérence de données répliquées. Je compte aussi travailler sur d'autres questions ouvertes tels que la structure du réseau et l'aspect sécurité.

**Cohérences des objets répliqués et interactions** Bien sûr, je continuerai à travailler sur les structures de données répliquées d'objets mutables, tels que les CRDT. Il serait intéressant d'adapter automatiquement le comportement des différents objets par rapport à leur rôle sémantique. Nous avons commencé par proposer plusieurs comportements lors d'un conflit. Par exemple, l'ajout d'un nœud pendant la suppression de son père. On peut aussi composer différents types de données répliquées ayant différents comportements. Ainsi on pourra composer des petites opérations en macro-opérations tout en gardant la commutativité.

Pascal Molli et Luis Ibanez travaillent sur la convergence d'un graph sémantique pour les objets répliqués. On pourrait relier ces travaux avec ceux que nous avons faits sur la composition et les vues. Mais d'un autre côté, on peut combiner nos travaux avec ceux d'Hala Skaf-Molli et Khaled Aslan sur l'émergence d'une ontologie pour certaines applications comme un gestionnaire de version distribué<sup>11</sup> afin d'y mesurer la divergence. Je propose d'utiliser ces informations pour proposer des optimisations du graphe de connexion entre les participants pour réduire «les temps» de la divergence. Je propose aussi d'aller dans ce sens pour les collaborations Man/Machine. Ces pistes devraient apporter des contributions au projet ANR KOLFLOW (<http://kolflow.univ-nantes.fr/>)

On pourrait adapter notre approche d'objet mutable à d'autres systèmes d'objets répliqués ayant d'autres contraintes. Par exemple, je pense qu'il serait possible de définir le même type d'objet sur les bases NoSQL (CouchDB, Riak,...). Actuellement ces bases de données distribuées fonctionnent sur des systèmes de map/reduce et de clef-valeurs. Ce système ne permet pas de compter de manière efficace, ni de faire des agrégations. Par exemple, la question : «combien de jouets possède paul» n'est pas traitée de manière efficace. Actuellement tout les réplicas répondent disant j'ai un jouet de paul et un noeu compte<sup>12</sup>. Le but serait de changer le système clefs-valeurs par un système complexe muni d'un grand nombre d'entités en interactions régies par des règles sémantiques. Ça impliquerait la possibilité de faire des cubes de données de manière décentralisée et efficace. C'est à dire qu'on pourrait faire toutes les agrégations possible au moment des insertions et de manière distribuée. Les métadonnées du web pourraient aider à générer automatiquement ce cube.

**Intégration dans le labex CominLabs** Le web social fait référence à un lieu de socialisation. Un lieu dont l'une de ses fonctions principale est de faire interagir beaucoup d'utilisateurs entre eux. Mes travaux s'inscrivent dans le domaine des systèmes collaboratifs à large échelle. Ils peuvent s'adapter sur des «cloud», ou dans certaines conditions sur des réseaux pair-à-pair. Ils peuvent même être utilisés dans des bases NoSql tel que dynamoDB d'Amazon, HBase utilisée par Facebook, ou autres... Ils participent au «web social» qui s'inscrit dans le cadre du Labex.

---

<sup>11</sup>Khaled Aslan, Hala Skaf-Molli, Pascal Molli, Connecting Distributed Version Control Systems Communities to Linked Open Data, CTS 2012

<sup>12</sup>Kai Orend, Florian Matthes, Thomas Büchner, Analysis and Classification of NoSQL Databases and Evaluation of their ability to Replace an Object-relational Persistence Layer

**Sécurité et intégrité** Un autre point intéressant des systèmes distribués est la sécurité des données. Cela peut aller de la non-divulgaration d'informations à la robustesse du réseau contre certaines attaques ou pannes.

L'équipe GDD est l'endroit parfait pour développer ces concepts avec : Patricia Serrano-Alvarado qui travaille sur la non-divulgaration de certaines données (vie privée) dans un réseau social, avec les données distribuées à l'aide d'une DHT. Il est possible de définir des politiques d'accès à différentes parties d'un document<sup>13</sup>. Il serait possible de structurer les données pour accepter les très grands documents. Le problème est que les opérations ou le mécanisme de synchronisation peuvent référer à des éléments «privés» pour être effectués correctement, ce qui impliquerait une fuite. Par exemple dans un arbre, une opération peut référer au chemin qui pourrait contenir des informations privées. Une idée serait de le remplacer les éléments secrets par des identifiants qui permettrait de garder la cohérence du document sans dévoiler le document complet. Le but serait d'exhiber des propriétés de non-interférence adaptée à ce cadre.

Yann Busnel travaille sur la détection d'attaque au niveau de la charge des pairs<sup>14</sup>. Ce type de détection est primordial dans un système à large échelle social. Une question reste ouverte : quel mécanisme peut-on mettre en place en cas de ce genre d'attaques ?

Achour Mostefaoui s'intéresse aux pannes byzantines sur ce type de réseaux<sup>15</sup>. On pourrait étudier l'impact sur des objets répliqués tels que les CRDT dans ces conditions. En admettant que les messages soient signés et que le Byzantin est identifié. Doit-on supprimer toutes ses contributions ? Nous devrions d'abord définir l'état le plus récent qui serait en quelque sorte valide. Il serait possible de définir des conditions afin de garantir la convergence des répliqués. Ceci n'est pas garanti même si le processus peut continuer.

**ANR STREAMS** Je pourrai aussi travailler avec Pascal Molli et Achour Mostefaoui sur l'ANR STREAMS(<http://streams.loria.fr/>). J'ai quelques idées concernant certains livrables, notamment le travail en sous-groupe avec la propriété d'isolation de chaque parti. La resynchronisation en cas de trop grande divergence qui pourra être issue des travaux portant sur les attaques byzantine. Nous pourrions renforcer la présence de l'équipe dans le projet.

**Conclusion** Mon insertion dans l'équipe GDD viendrait renforcer les compétences dans le domaine de la réplication optimiste de données dans les systèmes répartis pair-à-pair. Mon apport se distingue par ma capacité à faire le pont entre la modélisation formelle de mes propositions et leurs validations pratiques sous forme de prototypes logiciels. En effet, dans l'ensemble des travaux que j'ai menés jusqu'à présent, j'ai à la fois modélisé formellement mes propositions, démontré des résultats d'impossibilités ou leurs corrections, réalisé des implémentations et validé de manière pratique ces prototypes logiciels. Issu d'une équipe d'informatique fondamentale orientée vers la modélisation et la vérification, je possède un solide bagage dans le domaine. D'un autre côté, je conçois et développe des logiciels depuis longtemps.

## 2.5 Publications

---

Les liens vers les éditeurs sont sur ma page web <http://www.loria.fr/~stemarti/recherches> Les trois papiers publiés avec celui soumis à PODC'12 seront fournis en version papier si je suis convoqué à l'audition.

---

<sup>13</sup> Asma Cherif, Abdessamad Imine, Michaël Rusinowitch: Optimistic access control for distributed collaborative editors. SAC 2011: 861-868

& Asma Cherif, Abdessamad Imine, Michaël Rusinowitch: An Optimistic Mandatory Access Control Model for Distributed Collaborative Editors. Secure Data Management 2009: 89-106

<sup>14</sup> Emmanuelle Anceaume, Yann Busnel, Sebastien Gambs, AnKLe: Detecting Attacks in Large Scale Systems via Information Divergence, EDCC 2012

<sup>15</sup> Olivier Baldellon, Achour Mostefaoui, Michel Raynal: A Necessary and Sufficient Synchrony Condition for Solving Byzantine Consensus in Symmetric Networks. ICDCN 2011

### **2.5.1 Articles de conférences internationales avec comité de lecture**

- [1] Stéphane Martin, Pascal Urso, Stéphane Weiss, Scalable XML Collaborative Editing with Undo. OTM'10 Proceedings of the 2010 international conference on On the move to meaningful internet systems coopis 2010 (Short Paper): 507-514. (Approx. 20% des papiers acceptés)
- [2] Stéphane Martin, Denis Lugiez : Fixing Collaborative Editing on Typed Documents. Cooperative Design, Visualization, And Engineering 2010 : 19-26. (Approx. 25% des papiers acceptés)
- [3] Stéphane Martin, Denis Lugiez: Collaborative peer to peer edition: Avoiding conflicts is better than solving conflicts. IADIS Applied Computing (2) 2009: 124-128 (Approx. 20% des papiers acceptés)

### **2.5.2 Livrable ANR**

- [4] Stéphane Martin, Mehdi Ahmed-Nacer , Pascal Urso, Abstract unordered and ordered trees CRDT. <http://hal.inria.fr/hal-00648106/fr/>

### **2.5.3 Communication nationale sans actes**

- [5] Stéphane Martin, Denis Lugiez, Collaborative peer to peer edition: Avoiding conflicts is better than solving conflicts, ANR Codex (2009) <http://codex.saclay.inria.fr>.
- [6] Stéphane Martin, Denis Lugiez, Poster : Peer to peer collaborative editing on XML-like tree. Workshop CDUR'08, NOTERE.

### 3.1 Synthèse des enseignements effectués

L'ensemble des enseignements que j'ai effectués en qualité de moniteur à l'université de Provence Marseille (2007-2011) et en qualité d'Attaché Temporaire d'Enseignement et de Recherche à l'université de Provence (2010-2011), à l'université Henri Poincaré (2011-2012) est résumé dans le tableau suivant. Les durées indiquées sont en heure « équivalent TD ». Une description détaillée du contenu de chacun de ces enseignements est donnée dans la section 3.4.

TAB. 3.1: Synthèse des enseignements effectués - un total 514 heures effectif (Camembert) ou 558 heures Equivalent TD.

Période (s)	Enseignement	Type	Niveau	Horaire
2011-2012 ATER	Système de gestion de base de données	TP	L3	10 h
	Réseaux	Cours/TD/TP	L3	32 h
	Web	TP	L3	12 h+20 h
	Programmation Application Réparties	TP	Esial <sup>1</sup> Année 3	12 h
	RABD <sup>2</sup>	TP	Esial Année 3	10 h
	Architecture des ordinateurs	Cours/TD/TP	IUT Année 1	52 h
	Initiation au java	TD/TP	IUT Année 1	40 h
2010-2011 <sup>3</sup> ATER	Initiation à l'informatique	Cours/TD/TP	L1 Math info.	72 h
	Initiation à l'informatique	TP	L1 Bio.	20 h
	Architecture des ordinateurs	TD/TP	L3 Info.	40 h
	Systèmes d'exploitation	TD/TP	L3 Info.	40 h
2009-2010 Moniteur	Programmation Parallèle et Distribuée	TD/TP	M1 Info.	40 h
	Stage C	TP	L3 info.	30 h
2007-2009 Moniteur	Introduction à la programmation java	Cours/TP	M2 GSI <sup>4</sup>	2×64 h

### 3.2 Expériences et visions

Depuis 2007, j'ai pu enseigner à différents niveaux, à différents publics et dans différentes structures (figure 3.1). Je considère que la mission de la transmission du savoir est aussi primordiale que la recherche. Pour chacun de mes enseignements, j'essaie d'adapter le contenu par rapport aux étudiants. J'essaie de trouver des exemples et exercices intéressants pour étudiants, par exemple, en illustrant le problème purement calculatoire avec une histoire réaliste. Par exemple, j'ai illustré les problèmes de débits et de latences en faisant une course entre un chien (Woolfy) possédant une clef USB et une liaison ADSL. On m'a posé la question sur l'utilisation des satellites, j'ai donc fait un exercice dessus en

<sup>1</sup>Ecole Supérieure d'Informatique et Applications de Lorraine

<sup>2</sup>Répartition & Administration des Bases de Données

<sup>3</sup>Année basée sur 11 mois au lieu de 12

<sup>4</sup>Génie Statistique et Informatique



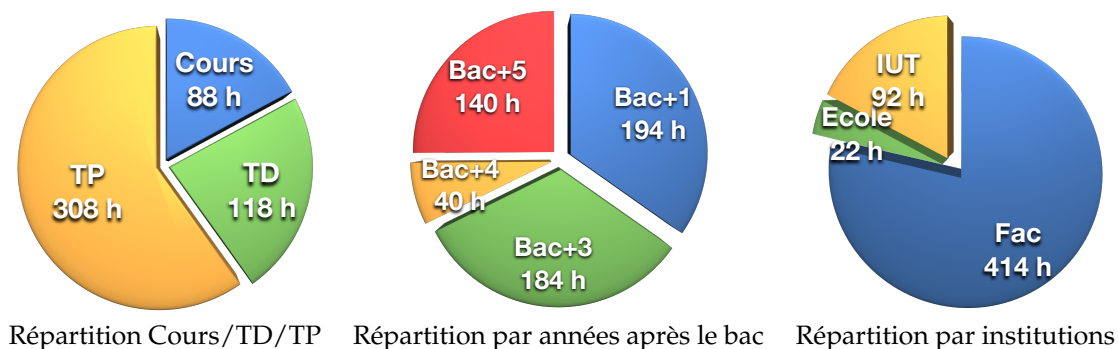


FIG. 3.1: Statistiques d'enseignements

le comparant avec le câble en fibre optique transatlantique et le pauvre Woolfy. J'essaie de proposer des projets attrayants, sans m'éloigner de l'objectif pédagogique tout en essayant de sortir du cadre scolaire : comment appliquer le contenu du cours sur des choses qui existent et qui sont susceptibles de les intéresser (jeux du plombier, match de foot, irc, serveur web, ...).

J'ai commencé l'enseignement au début de ma thèse avec des Cours/TP de java pour les Master Genie Statistique et Informatique. L'encadrement de ce petit groupe orienté mathématique m'a permis d'avoir une vision différente du domaine et de prendre confiance en moi. Il m'a permis de voir les avancés de chacun ce qui m'a conforté dans l'envie d'enseigner. J'ai assuré cet enseignement deux années de suite avant de m'essayer à d'autres modules.

Depuis, je suis parfaitement à l'aise avec tous les niveaux d'enseignement du L1 au M2 et avec tout type d'enseignement (TP/TD/CM)

Cette année, je découvre différentes institutions telle que l'IUT ou l'école d'ingénieur. J'ai acquis d'autres manières d'enseigner. Les attentes pédagogiques des étudiants dans ces établissements sont différentes. Le rythme, les temps d'attention et la capacité d'abstraction des étudiants le sont aussi.

Les difficultés d'enseigner en IUT sont dues au manque d'autonomie des étudiants et la nécessité d'être plus attentif au comportement, ne rien laisser passer afin d'instaurer un calme. De plus dans cet IUT, j'ai pris en charge le module «architecture des ordinateurs et systèmes d'exploitation» de SRC où tous les étudiants ne venaient pas de filière scientifique. Ceci m'a obligé à rédiger des exercices beaucoup moins abstraits faisant référence à la vie de tous les jours. J'ai dû aussi reprendre les bases sans que ceux venant d'une filière scientifique s'ennuient. Par exemple : faire des exercices bonus, démonter une machine, piéger les étudiants avec un programme malicieux ressemblant à une image, etc.

Malgré cela, j'ai adoré ce type d'enseignement pour la proximité, le retour (on sait de suite si le cours est intéressant) et la possibilité de suivre les progrès des étudiants. Bref, c'est une autre façon d'enseigner.

J'ai déjà été responsable de modules comme «architecture des ordinateurs et systèmes d'exploitation au DUT SRC» et «Initiation des réseaux en L3 informatique». Dans d'autres cas, j'ai à de nombreuses reprises rédigé des planches d'exercices de TD, TP, projet et aussi d'examens. J'ai aussi à de nombreuses reprises fait l'évaluation des projets d'un groupe.

L'enseignement me fait voir certains concepts sous un autre angle, ce qui me donne, étrangement, des idées dans ma recherche.

### 3.3 Projet d'enseignement et intégration

Mes activités d'enseignements s'articulent principalement autour de deux axes, à savoir : i) Systèmes distribués, programmation Web d'une part et ii) Architecture, systèmes d'exploitation réseaux, systèmes embarqués d'autre part.

#### 3.3.1 Systèmes distribués et répartis

Cette partie est dans mon domaine de recherche au sens large : les systèmes distribués et répartis.



J'ai auparavant participé à un module de programmation parallèle et distribuée enseigné en master. J'ai écrit les TD, les TP, le projet et la moitié des examens pour une trentaine d'étudiants. Nous avons vu les différents problèmes de ressources, de cohérence, horloge, verrous, etc. Nous avons illustré les problèmes fondamentaux du cours en TP à l'aide de Java RMI. Mes travaux de recherche me permettent de donner aux étudiants des exemples concrets pour appliquer directement le cours. Par exemple, il est possible de prendre un t'chat pour illustrer les horloges logiques. Le projet se basait sur un match de foot entre agents mobiles. Les agents n'étaient pas tous distribués au départ, ils n'avaient pas tous le même comportement.

J'ai découvert les EJB dans les TP de Programmation d'Applications Réparties. Dans mes TP d'outil Web, j'ai enseigné des composants de Java EE comme les JSP, JSF et JPA.

Je peux donc assurer des enseignements autour des systèmes distribués (partages et gestion des ressources, cohérence de données répliquées, découverte de services, exécution de code distant,...). J'ai enseigné quelques composants de Java EE (RMI, EJB, JNDI, JSF,...) et je connais SOAP/WSDL. J'ai appris à utiliser cette technologie pendant mon cursus et durant mon stage de Master. Je suis prêt à prendre en charge les enseignements d'**Outils pour les applications réparties**(Master 2).

Je peux aussi enseigner les **Applications et Technologies Web** en Master 1.

### 3.3.2 Architecture, réseaux et systèmes embarqués

**Architecture des ordinateurs & réseaux** En plus de mes activités de recherches, et depuis un certain temps, je m'intéresse particulièrement à l'électronique, la programmation de systèmes embarqués et les réseaux. Ce qui me permet de m'investir dans l'enseignement des matières comme : Architecture des ordinateurs, Systèmes d'exploitation, Réseaux, Programmation C ou assembleur, ou programmation synchrone et temps réel. En effet, j'ai été responsable de l'unité d'enseignement «architecture des ordinateurs et systèmes d'exploitation» pour le DUT SRC (60 étudiants) et du réseau pour la L3 informatique (15 étudiants).

C'est donc dans la continuité que je propose de prendre en charge l'UE de L3 et M1 **Réseaux informatiques**. Je suis prêt à m'investir dans les modules **Réseaux Haut Débit** et **Réseaux mobiles et sécurité**.

Je me propose aussi pour le cours d'**architecture des ordinateurs en Licence 3**. Je l'ai déjà enseigné.

**Systèmes embarqués & méthodes formelles** Mon intérêt pour les petits systèmes ne se limite pas au principe de fonctionnement. Je m'intéresse aux systèmes embarqués et à leur aspect formel. Je pense qu'il est important de sensibiliser les étudiants aux problèmes de petits systèmes qui les entourent. Sans le savoir, ils parient leurs vies sur leur calcul et leur réactivité (ex: l'ABS d'une voiture, systèmes médicaux,...)

Je connais les prouveurs UPPAAL (Automate/CTL) et spin (promella/LTL) pour les avoir étudiés en Master. Je me suis servi d'UPPAAL à plusieurs reprises pour illustrer le cours. Par exemple, pour vérifier un algorithme proposé par les étudiants. Le côté visuel du logiciel a permis à certains étudiants de comprendre les enchevêtrements qui étaient encore trop abstraits pour eux.

Je me propose pour l'enseignement du module **Modélisation et vérification des processus communicants** en Master 2. Ce module est à la fois assez proche de mes intérêts pour les systèmes mobiles et leur formalisation et à la fois proche de mon domaine de recherche (Systèmes répartis et concurrents).

Je suis aussi ouvert à d'autres unités d'enseignement.

### 3.3.3 Autres enseignements

Je suis ouvert à d'autres types d'enseignements plus «généraux».

**Algorithmique et programmation** Je peux aussi m'investir dans d'autres matières comme l'algorithmique de base, cours de programmation-objet (Java, C++, C#). J'ai enseigné à plusieurs reprises avec différents publics (L1 bio, Iut Info et L1 Info) l'algorithmique de base. J'ai aussi enseigné la programmation-objet au Master GSI. Le langage Java revient souvent dans différents enseignements comme les applications web, programmation des applications réparties, etc.

**Génie logiciel** Même si je ne l'ai pas encore enseigné, je suis initié aux techniques «modernes» de développement, de gestion et de maintenance de projets.

J'utilise depuis que je suis étudiant des forges ne serait-ce que pour le regroupement d'outils pratiques. L'université n'en proposant pas, j'en ai installé à la maison. Mes recherches durant ma thèse portant sur l'édition collaborative m'ont permis d'approfondir les mécanismes de ce genre d'outils collaboratifs. J'ai développé des prototypes et articles en utilisant SVN, So6 et GIT. Je connais les annotations de codes, sa documentation (javadoc/doxygen) et les tests unitaires (junit/Cunit/Sanity) pour les avoir utilisés.

J'utilise souvent des parties de UML pour le développement de certains projets (le diagramme de séquence, de classes,...). Je peux expliquer les liaisons entre les différents diagrammes. La recherche fait qu'on applique la méthode agile. Je connais les diagrammes de Gantt/pert pour l'allocation des ressources et l'estimation du temps. Il y en a quelques-uns dans les propositions de projets.

Je suis prêt pour enseigner les différentes méthodes pour ériger et suivre un projet logiciel. Je suis prêt à prendre en charge progressivement des modules tel que **Génie logiciel à objets** ou **Gestion de projets informatiques & gestion des travaux collaboratifs** tous deux en Master 1.

## 3.4 Détail des enseignements

---

### 3.4.1 ATER (2011-2012)

#### Cours/TD/TP de Réseaux (32h).

*Responsable* : Stéphane Martin  
*Effectif* : 15 étudiants  
*Travail fourni* : Cours, écritures des planches TD/TP et encadrement  
*Niveau* : Troisième année de licence d'informatique

Cette matière est une introduction aux réseaux. La difficulté est d'exposer suffisamment d'éléments pour intéresser les étudiants sans trop parler des éléments au programme du Master.

Les sujets abordés :

- Média de communication (Paire torsadée, Laser, Onde radio, etc ...)
- Notion de débit
- Différentes couches OSI
- Différentes couches de la pile TCP/IP
- Adresses et Masques
- API en C
- DNS.

#### TP de Programmation Application Réparties (12h).

*Responsable* : François Charoy <[francois.charoy@loria.fr](mailto:francois.charoy@loria.fr)>  
*Effectif* : 20 étudiants  
*Travail fourni* : Encadrement TP  
*Niveau* : Troisième année d'école d'ingénieur

Ces TP ont pour but l'utilisation des notions vues en cours avec différentes interfaces de programmation d'appels à distance, tels que le RMI et les EJB.

#### TP de Répartition & Administration des Bases de Données (10h)

*Responsable* : Nacer Boudjlida <[Nacer.Boudjlida@loria.fr](mailto:Nacer.Boudjlida@loria.fr)>  
*Effectif* : 20 étudiants  
*Travail fourni* : Encadrement TP  
*Niveau* : Troisième année d'école d'ingénieur

Ces TP s'axent principalement sur l'administration de bases de données Oracle. Les étudiants doivent configurer de A à Z tout le système Oracle sur une machine virtuelle.

### TP de Système de Gestion de Base de Données(10h)

*Responsable* : Nacer Boudjlida <Nacer.Boudjlida@loria.fr>  
*Effectif* : 15 étudiants  
*Travail fourni* : Encadrement TP  
*Niveau* : Troisième année de licence d'informatique

Ces TP porte sur l'application en SQL de la modélisation relationnel de divers problème vu en TD.

### TP d'Application Web (12h+20h)

*Responsable* : Pascal Urso <pascal.urso@loria.fr>  
*Effectif* : 15 étudiants  
*Travail fourni* : encadrement TP  
*Niveau* : Troisième année de licence d'informatique

Ces TP ont pour but de faire découvrir aux étudiants de L3 informatique, quelques concepts d'application web (architecture, session, ...) à l'aide des interfaces de programmation Java EE, telle que : les JSP, JSF et JPA.

### Architecture des ordinateurs (52h)

*Responsable* : Stéphane Martin  
*Effectif* : 59 étudiants  
*Travail fourni* : Cours (8h), Planches d'exercices TD et TP, encadrement TD/TP (16h/24h), évaluation  
*Niveau* : première année de DUT SRC

Ces cours TD et TP s'adressent aux étudiants de première année du département Services et Réseaux de Communication de l'IUT de St-Dié. Le but de cette matière est d'expliquer le principe de fonctionnement d'un ordinateur sans trop entrer dans les détails. Elle répondra notamment aux questions suivantes :

- Comment fait-on un calcul ?
- Comment stocker de l'information ?
- Quel sont les composants présents dans l'unité centrale ?
- Qu'est-ce qu'un système d'exploitation ? A quoi sert-il ?
- Qu'est-ce qu'un processus ?
- Qu'est-ce qu'une arborescence de fichiers ?
- Qu'est-ce qu'un droit d'accès ?
- Comment peut-ont partager un fichier ?
- ...

### TD/TP d'initiation au Java (40h)

*Responsable* : Abdellatif Bourjij <abdellatif.bourjij@iutsd.uhp-nancy.fr>  
*Effectif* : 50 étudiants  
*Travail fourni* : Encadrement TD/TP  
*Niveau* : première année de DUT Informatique

Le but de ce TD/TP est d'initier les étudiants de première année du département informatique de l'IUT de St-Dié au langage de programmation Java. Ils apprennent dans cette matière les structures de bases, les conditions, les boucles et les procédures avant de finir dans le paradigme de la programmation objet.

### 3.4.2 ATER (2010-2011)

#### Cours/TD/TP d'initiation à l'informatique (20h+22h+20h=72h)

*Responsable* : Pierre-Alain Reynier <[pierre-alain.reynier@lif.univ-mrs.fr](mailto:pierre-alain.reynier@lif.univ-mrs.fr)>  
*Effectif* : 20 étudiants  
*Travail fourni* : Cours, encadrement TD/TP, et évaluation  
*Niveau* : Première année de licence de mathématique et informatique

Cette matière a pour but de faire découvrir ce qui se cache derrière le mot informatique :

- Rôles des différents composants : Systèmes d'exploitation, interface (graphique et shell), processeur, bus,...
- Codage de l'information, calcul binaire et conversion.
- Shell (lignes de commandes).
- Algorithmique (python).
- Un petit peu d'architecture et de circuits logiques.

#### TP initiation informatique (20h)

*Responsable* : Mélanie Pomarès <[Melanie.Pomares@oamp.fr](mailto:Melanie.Pomares@oamp.fr)>  
*Effectif* : 20 étudiants  
*Travail fourni* : encadrement TD/TP, et évaluation  
*Niveau* : Première année de licence biologie

La thématique est la même que la précédente. Mais le langage de programmation étudié est *perl*.

#### TD/TP Architecture des ordinateurs(20h+20h)

*Responsable* : Peter Niebert <[peter.niebert@lif.univ-mrs.fr](mailto:peter.niebert@lif.univ-mrs.fr)>  
*Effectif* : 30 étudiants  
*Travail fourni* : Encadrement des TD/TP, écriture des projets et évaluation des étudiants  
*Niveau* : Troisième année de licence en informatique

Cette matière expose les principes de fonctionnement des ordinateurs. Il s'agit de comprendre, à bas niveau, leur organisation :

- Circuits logiques combinatoires et séquentielles. Utilisation du simulateur TkGate.
- Programmation assembleur sur un microcontrôleur. Il s'agissait de lire et de comprendre la documentation d'une puce PIC16 afin de pouvoir réaliser un digicode. Des platines d'essais étaient disponibles.

#### TD/TP Système d'exploitation (20h+20h)

*Responsable* : Leonardo Brenner <[leonardo.brenner@univ-provence.fr](mailto:leonardo.brenner@univ-provence.fr)>  
*Effectif* : 35 étudiants  
*Travail fourni* : Encadrement TD/TP, Evaluation des étudiants  
*Niveau* : Troisième année de licence d'informatique

Cette matière a pour but d'entrer au coeur d'un système d'exploitation de type UNIX :

- Ordonnancement de processus
- Duplication de processus (fork)
- Signaux
- Systèmes de fichiers (ext, fat, ...)
- Gestion de la mémoire (pagination et segmentation)
- Exclusion mutuelle.

### 3.4.3 Monitorat (2007-2010)

#### Stage C (30h)

*Responsable* : Luigi Santoncanale <[lsantoca@cmi.univ-mrs.fr](mailto:lsantoca@cmi.univ-mrs.fr)>  
*Effectif* : 15 étudiants  
*Travail fourni* : Encadrement TP et évaluation  
*Niveau* : Troisième année de licence d'informatique

Introduction au langage C, et aux Makefiles.

#### TD/TP Programmation Parallèle et Distribuée (20h+20h)

*Responsable* : Rémi Morrin <[remi.morin@lif.univ-mrs.fr](mailto:remi.morin@lif.univ-mrs.fr)>  
*Effectif* : 30 étudiants.  
*Travail fourni* : Encadrement TD/TP, Planches de TD et de TP, évaluation  
*Niveau* : Première année de master informatique

- Introduction aux fondements des systèmes parallèles, concurrents, répartis et distribués: modèles, problèmes et algorithmes fondamentaux, formalisation.
- Programmation d'applications illustrant en pratique les paradigmes étudiés: utilisation d'API de haut niveau. (Java 1.5)

#### Cours/TP Java (64h 2007-2009)

*Responsable* : Stéphane Martin  
*effectifs* : 10 étudiants.  
*Travail fourni* : Encadrement Cours/TP, Planches de TP et évaluation  
*Niveau* : Deuxième année de master génie statistique et informatique

Introduction aux notions de programmation objet en Java.

Le principe était de donner les méthode de pour la modélisation orienté objet et de l'appliquer directement en Java.





## Charges Collectives

**Relecture d'article** : J'ai relu un article pour la conférence ACM Group 2012.

**Organisation des séminaires des doctorants 2009-2010** : Le but était d'organiser des séminaires entre doctorants de différents laboratoires. Cela permet à chacun de s'entraîner devant une foule et de faire un exposé clair à un public non initié, mais scientifique. Ce séminaire avait lieu une à deux fois tous les deux mois. Le public était composé des doctorants du LIF (informatique) et du LAMP (mathématiques). L'ambiance était assez détendue ce qui permettait facilement les questions, les conseils, etc.

**Diffusion de la science 2007-2010** : Pendant le projet CIES, nous avons essayé de développer l'esprit critique (Zététique) à des jeunes, allant du CM2 au collège, par des expériences ludiques. Nous avons aussi tenu un stand à la fête de la science. Cette expérience m'a appris à vulgariser la science et a permis de m'exercer à la communication vers un public non scientifique. Sous l'impulsion de Denis Caroti, «Association Marseille Zététique» une association loi 1901 a éclos à la suite de ce projet.

**Représentant des doctorants auprès de l'école doctorale 2007-2011** : Le but est de remonter les craintes ou les questions des doctorants à l'école doctorale et d'un autre côté diffuser l'information aux doctorants. Il est aussi d'assister aux conseils de l'école doctorale qui a lieu une fois par trimestre.

**Vice président d'une association étudiante 2006-2009** : Vice-Président d'une association loi 1901 d'étudiants, appelée le cyberFoyer. Son but est d'offrir aux étudiants du café, des friandises, viennoiseries et un endroit pour étudier avec des ordinateurs en libre accès.

Cela m'a permis d'entrevoir les difficultés d'encadrer une équipe de bénévoles, la gestion d'une salle et l'organisation d'événements pour les différentes promotions.







## Recommandations professionnelles

Des lettres de recommandations ont été obtenues auprès des personnalités scientifiques suivantes. Ces personnes peuvent être contactées pour de plus amples informations. Une copie de ces lettres est en section A.1.

- **Denis Lugiez** Professeur à Aix-Marseille Université  
Responsable de l'équipe Modélisation et Vérification au Laboratoire d'informatique Fondamentale  
Laboratoire d'Informatique Fondamentale  
CMI, 39, rue Joliot-Curie  
F-13453 Marseille Cedex 13, France.  
Tel: +33 (0)4 91 11 36 23;  
Fax: +33 (0)4 91 11 36 02;  
Tel (secrétariat): +33 (0)4 91 11 35 21  
[denis.lugiez@lif.univ-mrs.fr](mailto:denis.lugiez@lif.univ-mrs.fr)
  
- **Leonardo Brenner** Maître de conférences à Aix-Marseille Université  
Laboratoire d'Informatique Fondamentale  
CMI, 39, rue Joliot-Curie  
F-13453 Marseille Cedex 13, France.  
Tel : +33 (0)4 13 55 13 93  
Fax : +33 (0)4 91 05 60 33  
[leonardo.brenner@univ-provence.fr](mailto:leonardo.brenner@univ-provence.fr)
  
- **Pierre-Alain Reynier** Maître de conférences à Aix-Marseille Université  
Laboratoire d'Informatique Fondamentale  
CMI, 39, rue Joliot-Curie  
F-13453 Marseille cedex 13, France  
Tel : +33 (0)4 13 55 13 07  
Fax : +33 (0)4 13 55 13 02  
[pierre-alain.reynier@lif.univ-mrs.fr](mailto:pierre-alain.reynier@lif.univ-mrs.fr)
  
- **Pascal Urso** Maître de conférences à l'Université de Lorraine  
LORIA Campus Scientifique,  
Université de Lorraine  
54506 Vandoeuvre-lès-Nancy, Cedex, France  
Tel : +33 (0)3 54 95 86 29  
Fax : +33 (0)3 83 41 30 79  
[pascal.urso@loria.fr](mailto:pascal.urso@loria.fr)
  
- **Jean-Luc Husson** Maître de conférences à l'Université de Lorraine  
Responsable du département «Service et Réseaux de Communication» et directeur adjoint de l'IUT de Saint dié des Vosges  
IUT de Saint dié des Vosges  
11, rue de l'université

88100 Sain Dié des Vosges, France  
[jean-luc.husson@univ-lorraine.fr](mailto:jean-luc.husson@univ-lorraine.fr)

- **Odile Mella** Maître de conférences à l'Université de Lorraine  
LORIA Campus Scientifique,  
Université de Lorraine  
54506 Vandoeuvre-lès-Nancy, Cedex, France  
Tel : +33 (0)3 29 53 60 30  
fax : +33 (0)3 29 53 60 13  
[odile.mella@loria.fr](mailto:odile.mella@loria.fr)



## A.1 Copies des lettres de recommandations

---



LIF - UMR 6166  
CNRS

### Avis sur la candidature de Stéphane Martin sur un poste de Maître de conférences

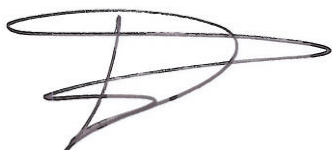
J'ai encadré Stéphane Martin pour sa thèse qui portait sur l'édition collaborative de documents semi-structurés. Ce sujet est différent de celui de son stage de Master 2 qui portait sur la vérification de protocoles cryptographiques et cela a été un handicap pour le début de sa thèse. Malgré cela, il s'est impliqué totalement dans son sujet et a su comprendre les questions qui se posent dans un domaine qui n'est pas toujours très formalisé et demande des connaissances assez larges en algorithmique, structures de données, calcul distribué, et modélisation. L'édition collaborative et la réalisation d'éditeurs collaboratifs est un sujet de recherche actif sur lequel travaillent des équipes tant en France qu'à l'étranger. Les applications sont nombreuses, depuis les logiciels couramment utilisés comme cvs et svn jusqu'aux utilisations dans le cadre de réseaux sociaux et de systèmes pair à pair. La question étudiée est pratique mais demande une formalisation et une approche rigoureuse pour être traitée correctement. Le problème est simple et d'actualité : comment éditer en pair à pair par des sites distincts des documents complexes en garantissant la convergence (chaque site a le même document à la fin du processus). Le point de départ a été la thèse d'Imine au LORIA, dont une partie des résultats est de montrer que la plupart des algorithmes proposés pour les structures séquentielles, i.e. les documents textuels, sont faux. La thèse de Stéphane s'intéresse aux structures plus riches représentant les documents XML qui sont devenus le format d'échange de document par excellence. Il a proposé plusieurs contributions très intéressantes que les spécialistes du domaine ont particulièrement appréciées lors de sa soutenance, aussi bien dans le cadre des transformées opérationnelles que dans le cadre CRDT (structures de données commutatives). Dans le premier cas, il a donné des jeux d'opérations permettant d'effectuer l'édition collaborative pair à pair (donc non centralisée, à la différence de svn par exemple) sur des données arborescentes. Il a aussi donné des résultats d'impossibilité d'existence de transformée opérationnelle pour certains jeux d'opérations, ce qui n'avait jamais été fait. Il a également pris en compte l'opération de déplacement de sous-arbre, qui est usuellement ignorée car trop difficile à traiter dans ce cadre. Pour l'approche CRDT, il a proposé un algorithme générique d'édition collaborative fondé sur une notion de dépendance sémantique en reprenant les idées de Lamport. Il a également donné une variante permettant d'ajouter du typage aux documents (par exemple une DTD) et il a pareillement étudié plusieurs jeux d'opérations, dont le déplacement de sous-arbre. Un résultat très intéressant de la thèse montre qu'il suffit de vérifier une propriété simple du jeu d'opération pour pouvoir utiliser son algorithme générique, ce qui est

un progrès considérable. Par ailleurs la complexité de l'algorithme est bien meilleure que pour l'approche transformée opérationnelle utilisant l'algorithme de Ressel. Il a généralisé son travail à une structure de graphe permettant d'avoir un cadre général exprimant simplement le remplacement de sous-arbre. Par ailleurs il a également montré comment prendre en compte une structure de donnée XML dans sa totalité (ordre entre fils, attributs, ...) et il a réalisé un prototype permettant de valider expérimentalement l'approche commutative comme l'analyse de complexité le laissait présager. Dans tout son travail, il a fait preuve d'initiative et d'un vrai sens de chercheur, malgré une difficulté à rédiger qui l'a handicapé dans la rédaction de sa thèse.

Je considère que Stéphane Martin a fait preuve des capacités nécessaires qu'on attend d'un enseignant-chercheur en informatique pour la recherche. Il a par ailleurs donné toute satisfaction dans son enseignement tant en premier cycle (initiation en L1) qu'en second cycle (programmation parallèle et distribuée). Son expérience à l'université de Provence est importante puisqu'elle correspond à un monitorat suivi d'un ATER. Je pense que l'Université qui le recrutera bénéficiera d'un maître de conférences enthousiaste et dynamique qui renforcera son potentiel en recherche et enseignement.

Marseille, le 22 février 2012

Denis Lugiez  
Professeur d'Informatique  
Directeur de l'UFR Mathématiques-Informatique-Mécanique  
Université de Provence  
CMI, 39 r Joliot-Curie  
F-13453 Marseille Cedex 13  
denis.lugiez@lif.univ-mrs.fr



A qui de droit,

### **Lettre de recommandation**

Je soussigné, Leonardo Brenner, responsable de l'UE de Systèmes d'Exploitation à l'Université de Provence, atteste que Stéphane Martin a fait partie de l'équipe pédagogique de cette UE.

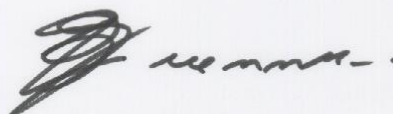
L'année de 2010 a été ma première année comme responsable de cette UE et où il avait un travail conséquent de rédaction, mise à jour à faire. Stéphane Martin s'est impliqué d'une façon conséquente dans le sujet, toujours cherchant à approfondir ces connaissances dans la matière pour mieux enseigner à ses étudiants.

Il a montré, durant les enseignements, beaucoup de professionnalisme et de sérieux. Il a une aisance remarquable en face des étudiants.

Pour toutes ces raisons, je le recommande chaudement à un poste de Maître de conférences.

À Marseille, le 13 Décembre 2011.

Leonardo Brenner  
Maître de Conférences  
Université de Provence  
Leonardo.Brenner@univ-provence.fr





## Lettre de recommandation pour

### **M. Stéphane MARTIN**

Moniteur puis ATER à l'Université de Provence, Stéphane Martin a effectué de nombreux enseignements dans les filières informatiques. Plus particulièrement, il est intervenu au cours de l'année universitaire 2010/2011 dans l'UE Introduction à l'Informatique en L1 Mathématiques et Informatique dont j'étais responsable. Dans le cadre du plan réussite en licence, cette unité d'enseignement est dispensée en groupes, et Stéphane Martin était chargé de l'ensemble des enseignements (Cours, Travaux dirigés, Travaux pratiques) pour un des groupes.

Ses enseignements m'ont pleinement satisfait, et j'ai pu constater son sérieux et son investissement. Stéphane Martin est très dynamique et a prouvé sa capacité à concevoir des sujets de TD ou de devoirs de façon autonome. Il a également pris l'initiative de créer des supports de cours sous la forme de transparents et a été apprécié des étudiants.

Par ailleurs, Stéphane Martin est un collègue disponible et agréable.

Pour toutes ces raisons, j'atteste ici des compétences de Stéphane Martin en matière d'enseignement, et suis convaincu que l'établissement qui le recrutera sera satisfait de son travail d'enseignant.

Marseille, le 23 mars 2012.

Pierre-Alain Reynier  
Maître de Conférences à l'Université d'Aix-Marseille

Adresse : LIF, CMI, 39 rue Joliot Curie, 13 453 Marseille Cedex 13  
Tél. : +33/0 4 13 55 13 07 / Fax : +33/0 4 13 55 13 02  
Courriel : pierre-alain.reynier@lif.univ-mrs.fr



**Pascal URSO**  
Maître de Conférences  
Equipe SCORE - LORIA  
Université de Lorraine  
pascal.urso@loria.fr

## Lettre de Recommandation pour M. Stéphane Martin

J'ai eu l'occasion de connaître M. Stéphane Martin, alors doctorant à l'Université de Provence, en 2009 lors d'un séminaire où il présentait ses travaux de recherches au Loria. Grâce à la convergence de nos thèmes de recherche, nous avons débuté une collaboration. Cette collaboration s'est révélée fructueuse notamment à travers la publication d'un article à la conférence CoopIS'2009. Après l'année 2010 consacrée à la rédaction de sa thèse, je lui ai suggéré de rejoindre l'équipe SCORE en tant qu'A.T.E.R. à temps complet en septembre 2011. Dès son arrivée, il s'est parfaitement intégré dans notre équipe et son expertise dans le domaine de l'édition collaborative de données structurées a permis d'obtenir rapidement (malgré une lourde charge d'enseignement) des résultats très intéressants. Ces travaux, qui ouvrent la voie à la réplication à large échelle de données structurées sans synchronisation ni consensus, sont déjà rédigés sous la forme d'un livrable pour l'ANR ConcoRDanT et ont été soumis à PODC'2012.

Je rajouterais que je lui ai confié, ainsi que d'autres permanents de l'équipe SCORE, des enseignements et que nous sommes très satisfaits de son implication et de son contact avec les étudiants.

Je soutiens donc fortement la candidature de M. Martin à la fonction de Maître de Conférences. Sa pertinence et son esprit scientifique sont d'un grand niveau, et sa collaboration avec l'équipe SCORE est fructueuse.

Bien cordialement,

Fait à Nancy, le 22 mars 2012



Nancy-Université



Campus scientifique  
BP 239 – 54506 Vandœuvre-lès-Nancy Cedex  
Tél. : 03 83 59 20 00 – Fax : 03 83 27 83 19 www.loria.fr



A qui de droit,

Stéphane Martin est actuellement sur un poste d'ATER comprenant des enseignements à l'IUT de Saint-Dié des Vosges et d'autres au Département Informatique de la Faculté des Sciences de Nancy. C'est dans ce dernier cadre, qu'il a accepté de prendre en charge, cet automne, l'intégralité de l'UE Introduction aux Réseaux et à la programmation réseau de la 3<sup>e</sup> année de la Licence d'Informatique.

Malgré le fait que cet enseignement faisait partie d'un semestre répété et que tous les documents pédagogiques lui étaient fournis, Stéphane Martin s'est fortement impliqué dans l'UE. Il a fait preuve de curiosité, d'initiative et d'autonomie en modifiant ces documents et en aménageant le contenu pédagogique de l'UE à sa façon. Ces qualités sont pour moi des qualités importantes pour le métier d'enseignant-chercheur, je recommande donc sa candidature à un poste de Maître de Conférences.

Nancy, le 23 février 2012,



Odile Mella  
Maître de Conférences à l'Université de Lorraine (ex Université Henri Poincaré de Nancy)  
Equipe Parole Loria  
03 83 59 20 80  
odile.mella@loria.fr



Saint-Dié-des-Vosges, le 05 mars 2012

**Objet :** Avis pédagogique sur M Stéphane Martin.

**Jean-Luc Husson**

**Maître de conférences Informatique**

Chef du département « Services et Réseaux de Communication »

Responsable des LP ATC « Activités et Techniques de Communication »

Directeur adjoint de l'IUT

IUT de Saint Dié des Vosges  
11, rue de l'Université  
88100 Saint Dié des Vosges

Tél : 03 29 53 60 30  
Fax : 03 29 53 60 13

[jean-luc.husson@iutsd.uhp-nancy.fr](mailto:jean-luc.husson@iutsd.uhp-nancy.fr)

M. Martin a intégré l'IUT de Saint-Dié-des-Vosges en septembre 2011 en tant qu'ATER 27<sup>ème</sup> section sur un ½ poste (96 H équivalent TD - HeqTD). Dans cette composante, son service d'enseignement a été réparti entre deux départements de l'IUT, le département Informatique et le département Services et Réseaux de Communication (SRC) que je coordonne.

Il est intervenu en première année du DUT SRC pour un total de 52HeqTD en assurant toutes les formes d'enseignement (8h CM, 16h TD et 24h TP) du module « Système d'exploitation et architecture des ordinateurs ». Il a également assuré l'évaluation des étudiants de ce module. Il faut noter que ce module est traditionnellement pris en charge par des vacataires ou des ATER, pour lesquels nous n'avons pas de référent permanent, et qui justifie la grande autonomie dont a dû faire preuve M Martin dans l'organisation de cet enseignement.

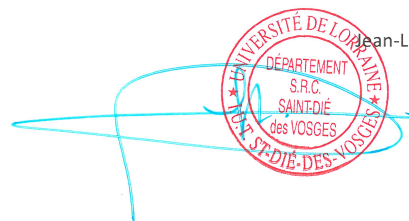
M. Martin a continuellement pris conseil auprès des enseignants de 27<sup>ème</sup> section pour bien saisir le contexte particulier de cet enseignement technique dans le diplôme très pluridisciplinaire du DUT SRC. Il a conjointement pris soin d'adapter ses enseignements au niveau de nos étudiants (premier cycle universitaire notamment) dont la diversité des origines (Bac ou diplôme B+2 pour les LP) est très forte.

Il faut noter que ce ½ service dans notre IUT a été complété par un ½ service dans l'UFR sciences à Nancy. La bilocalisation géographique induite (85 km de distance entre les deux centres), associée à son implication dans deux départements de l'IUT constituent un contexte de travail délicat. Cependant M Martin s'est montré disponible et souple dans l'organisation des emplois du temps. C'est pour ces raisons également que nous avons décidé de ne pas imposer à M Martin d'investissement administratif complémentaire tel que le suivi de stagiaires, la participation aux sessions de soutenances de projets tutorés et de stage ou à un forum d'orientation.

Compte tenu de notre expérience de quelques mois, nous avons pu remarquer les facultés d'adaptation de M Martin et pensons qu'il peut intégrer rapidement une équipe pédagogique. Aussi, nous formulons un avis tout à fait favorable à un recrutement sur un poste de maître de conférences dans les sections concernées par sa discipline de recherche.

Fait à Saint-Dié, le 05 mars 2012, pour servir et valoir ce que de droit.

Jean-Luc Husson





## A.2 Procès verbal de soutenance



### PROCES-VERBAL DE SOUTENANCE DE THESE

Le 08 Septembre 2011,

Etudiant : Monsieur STEPHANE MARTIN

Né le 07 Février 1984 à SAINT-DENIS (FRANCE)

Diplôme : Doctorat d'Informatique

Titre des travaux : Edition Collaborative des documents semi-structurés

Ecole doctorale : Mathématiques et informatique de Marseille (184)

Directeur : M. DENIS LUGIEZ

Codirecteur : Monsieur Inconnu

Lieu de soutenance : CMI Salle C101

La soutenance est publique

Résultat :

Mention :

Membres du jury

Monsieur RUSINOWITCH MICHAEL - directeur de recherche - INRIA

Monsieur MOLLIN PASCAL - professeur des universités - UNIVERSITE NANTES

Madame BENZAKEN VERONIQUE - professeur des universités - UNIVERSITE PARIS 11

Monsieur LUGIEZ DENIS - professeur des universités - UNIVERSITE AIX-MARSEILLE 1

Monsieur MORIN RÉMI - professeur des universités - UNIVERSITE AIX-MARSEILLE 2

Monsieur SHAPIRO MARC - directeur de recherche - UNIVERSITE PARIS 6

Signatures

NB : Il convient d'indiquer en marge de son nom quel est le Président du jury

Stephane Martin a exposé clairement ses contributions nombreuses et variées sur le thème de l'édition collaborative, un domaine actuellement en plein essor.

Dans le cadre de modèles pair à pair convergents (par transformée opérationnelle ou par type de données répliqué commutatif), le candidat présente des résultats novateurs, que ce soit des résultats d'impossibilité d'existence de certains types, ou que ce soit des résultats constructifs de nouveaux types de données, graphes ou des arbres répliqués convergents. L'idée d'utiliser des réparateurs de type avec une vue locale et une vue globale convergentes est également séduisante.

Le jury a apprécié la minutie et la rigueur de la vérification des protocoles proposés à l'aide d'un démonstrateur automatique mais ce manuellement. Stephane Martin a très bien répondu aux nombreuses questions, tant d'ordre général que technique, posées par le jury, et a ainsi prouvé sa maîtrise profonde de son sujet d'étude.

Page 1 / 1



### Avis du jury sur la diffusion d'une thèse soutenue

#### RENSEIGNEMENTS SUR L'AUTEUR

Nom patronymique : MARTIN

Prénom : STEPHANE

Nom marital : /

#### RENSEIGNEMENTS SUR LA THESE

Titre : Edition Collaborative des documents semi-structurés

Thèse sur travaux :  oui  non

Directeur de la thèse : DENIS LUGIEZ

Date de soutenance : 08 Septembre 2011

#### CONFIDENTIALITE DE LA THESE

Thèse confidentielle :  oui  non

Durée de confidentialité demandée :  2 ans  5 ans  10 ans  Autre : .....

#### AVIS DU JURY SUR LA DIFFUSION DE LA THESE

Thèses pouvant être diffusée en l'état

Thèse pouvant être diffusée après corrections demandées pendant la soutenance : un second dépôt de thèse est obligatoire.

Indiquez les corrections demandées au verso de la feuille.

Thèses ne pouvant être diffusée

Fait à Marseille....., le 8/09/2011.....

Le président du jury  
(Nom, prénom, signature)

MORIN Lémi



Contrat de mise en ligne d'une thèse soutenue

ETABLISSEMENT ET COMPOSANTE

Etablissement de soutenance : Université de Provence  
Faculté ou composante : MIM  
Etablissement en co-tutelle :  
Ecole doctorale : Mathématiques et informatique de Marseille (184)

RENSEIGNEMENTS SUR L'AUTEUR

NOM patronymique : MARTIN Prénom : STEPHANE  
Né(e) le : 07/02/1984 Né(e) le : Nationalité : FR  
Adresse électronique professionnelle : stephane.martin@etu.univ-provence.fr  
Adresse électronique personnelle :

RENSEIGNEMENTS SUR LA THESE

Titre : Edition Collaborative des documents semi-structurés  
Champ disciplinaire : Spécialité :  
Date de soutenance : ~~16/09/11~~ 08/09/2011

ATTESTATION DE CONFORMITE

Cochez les choix correspondants :

- L'auteur certifie la conformité de la version électronique à l'exemplaire imprimé déposé conjointement.  
 L'auteur ne certifie pas la conformité de la version électronique à l'exemplaire imprimé déposé conjointement.

AUTORISATION DE MISE EN LIGNE

Cochez les choix correspondants :

- L'auteur reconnaît avoir pris connaissance de la *Charte de dépôt et de diffusion électronique des thèses* d'Aix-Marseille Université en vigueur à la date de signature du présent contrat.  
 L'auteur garantit que tous les documents sont libres de droit ou qu'il a acquis les droits afférents pour la reproduction et la représentation sur tous supports.  
 L'auteur n'autorise pas la mise en ligne de la thèse.

- L'auteur autorise la mise en ligne de la thèse sans restriction sur intranet exclusivement (1)  
L'autorisation est ainsi consentie à l'Université pour une mise en ligne sur l'intranet de l'Université pour la durée légale de protection accordée par le droit de la propriété intellectuelle. Toutefois malgré les moyens mis en oeuvre, l'Université ne peut garantir à l'auteur un accès intégralement sécurisé sur intranet, elle ne pourra en conséquence être tenue responsable d'une diffusion au-delà de l'intranet.
- L'auteur autorise la mise en ligne de la thèse sans restriction sur internet (2)  
L'autorisation est ainsi consentie à l'Université pour le monde entier et pour la durée légale de protection accordée par le droit de la propriété intellectuelle.

Dispositions communes à tout type d'autorisation de diffusion :

L'auteur consent à ce que sa thèse fasse l'objet d'une mise à disposition gratuite du public concerné autorisé à la consulter, la fixer sur tout support et le représenter à titre gratuit pour un usage exclusivement personnel ou pédagogique.

L'auteur autorise la mise en ligne de la thèse par l'Université et l'accomplissement par celle-ci des actes de reproduction et de représentation nécessaires à cette fin sur tous supports. L'Université n'est en aucun cas autorisée à procéder à d'autres exploitations.

Il est expressément convenu que la présente autorisation est consentie à titre gratuit.

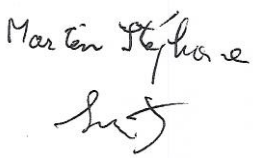
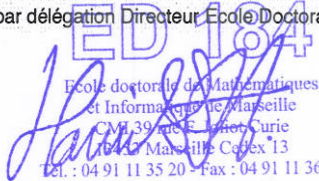
En conséquence, aucune rémunération ne sera versée par l'auteur en contrepartie de la mise en ligne de sa thèse.

L'autorisation est consentie à l'Université à titre non exclusif. L'auteur peut en conséquence exploiter librement sa thèse, sur laquelle il conserve l'ensemble de ses droits patrimoniaux, et peut notamment la faire publier par un éditeur de son choix, à l'exclusion des thèses confidentielles.

L'auteur peut retirer l'autorisation de mise en ligne à tout moment sans avoir à justifier de motif. Dans ce cas, il doit aviser l'Université de sa décision par lettre recommandée avec accusé de réception, selon la procédure indiquée par son université d'inscription. L'Université s'engage à retirer l'oeuvre mise en ligne au plus tard dans un délai de 3 mois à compter de la réception du LRAR.

La loi applicable au présent contrat est la loi française.

Le tribunal compétent pour juger de tout contentieux lié au présent contrat est le Tribunal Administratif de Marseille.

Fait à <u>MARSEILLE</u> ....., le <u>09/09/2011</u> .....	
L'auteur (Nom, prénom, signature)	Président de l'Université (ou par délégation Directeur Ecole Doctorale)
	 ED 184 Ecole doctorale de Mathématiques et Informatique de Marseille CM 39 des E. Henri Curie 10417 Marseille Cedex 13 Tel. : 04 91 11 35 20 - Fax : 04 91 11 36 47 ed184@cmi.univ-ms.fr

(1) Intranet : s'entend du réseau informatique accessible gratuitement depuis des postes individualisés mis à disposition des enseignants, des chercheurs, des étudiants et du personnel dans l'enceinte d'un établissement appartenant à l'Université de Provence et à distance après authentification sécurisée

(2) s'entend d'un réseau informatique mondial accessible au public sans identification préalable.





## A.3 Attestation d'enseignements

---



LIF - UMR 6166  
CNRS

### Attestation d'enseignement pour Stéphane Martin

Je certifie que monsieur Stéphane Martin a assuré les enseignements suivants comme moniteur et ATER à l'université de Provence.

- 2010-2011
  - Licence de Biologie L1. Initiation à l'Informatique (TP) 20h.
  - Licence Mathématiques Informatique L1. Introduction à l'informatique (cours + TD + TP) 72h.
  - Licence Informatique L3. Architecture des ordinateurs (TD + TP) 40h.
  - Licence Informatique L3. Systèmes d'exploitations L3 (TD + TP) 40h
- 2009-2010.
  - Licence Informatique L3. Stage C (TP) 30h
  - Master Informatique. Programmation parallèle et distribuée (TD + TP) 40h
- 2007-2009.
  - Master Génie Informatique et Statistique M2. Programmation Java. (TP) 64h/an

Marseille, le 6 décembre 2011

Denis Lugiez  
Professeur d'Informatique  
Directeur de l'UFR Mathématiques-Informatique-Mécanique  
Université de Provence  
CMI, 39 r Juliot-Curie  
F-13453 Marseille Cedex 13  
denis.lugiez@lif.univ-mrs.fr

Denis LUGIEZ  
Directeur de l'UFR-MIM



## A.4 Attestation directeur de l'école doctorale (Charges collectives)

### ECOLE DOCTORALE EN MATHÉMATIQUES ET INFORMATIQUE DE MARSEILLE

Centre de Mathématiques et d'Informatique  
Université de Provence  
39, rue Joliot Curie - F-13453 Marseille cédex 13  
tél. (33) 04 91 11 35 20      ed184@cmi.univ-mrs.fr      fax (33) 04 91 11 35 02

#### ATTESTATION

Je soussigné, Hamish SHORT, Directeur de l'Ecole Doctorale de Mathématiques et Informatique de Marseille, atteste que :

**Monsieur MARTIN Stéphane**

A rempli avec sérieux tout au long de sa thèse sa fonction de représentant des étudiants et s'est beaucoup investi dans les organisations de séminaire pour les doctorants.

Fait à Marseille, le 12 décembre 2011

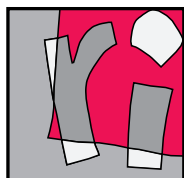
Hamish SHORT  
Directeur de l'Ecole Doctorale 184

Ecole Doctorale de Mathématiques  
et Informatique de Marseille  
CMI 39 rue F. Joliot Curie  
13453 Marseille Cedex 13  
Tél. : 04 91 11 35 20 - Fax : 04 91 11 36 47  
ed184@cmi.univ-ms.fr



## A.5 Rapports de présoutenance de Véronique Benzaken

---



LABORATOIRE DE RECHERCHE EN  
INFORMATIQUE

L.R.I - PCRI  
Bat. 650  
Université Paris Sud  
91405 Orsay Cedex

Tél. 01.69.15.66.28  
Fax : 01.69.15.66.13

Orsay, le 12 juillet 2011

Pr., Véronique . Benzaken  
LRI, UMR 8623, C.N.R.S  
Equipe ProVal  
Université de Paris Sud  
email : veronique.benzaken@lri.fr

---

### Rapport sur "Edition collaborative des documents semi-structurés" en vue de l'obtention du Doctorat en Informatique de l'Université Aix-Marseille, manuscrit présenté par Monsieur Stéphane Martin

**Contexte** Le développement constant d'internet a suscité des évolutions dans nos pratiques de travail. Il est désormais commun de collaborer "en réseau" afin de développer du code ou de rédiger un document commun. De nombreux outils tels que CVS ou Subversion (SVN) existent qui permettent la réalisation de ce type de tâches.

Parallèlement, le format XML s'est imposé depuis une dizaine d'années comme standard pour l'échange et la représentation de données sur internet. Ainsi, dans un contexte hautement distribué, il devient usuel que différentes applications aient accès à un même document. Par exemple les pages de Wikipédia sont régulièrement modifiées par des contributeurs de tous horizons.

Ce type d'applications est généralement qualifié d'édition collaborative. La difficulté majeure est de pouvoir garantir que chaque participant à l'élaboration du document ait, tôt ou tard, une version qui soit cohérente, c'est à dire identique à celles des autres. L'algorithme d'orchestration du tout doit donc garantir une propriété de *convergence*. Suivant le contexte, centralisé ou pleinement distribué, un tel processus de coordination s'avère profondément affecté.

Le contexte de travail de cette thèse est celui là. S. Martin se propose d'étudier la mise à jour collaborative de documents XML dans un environnement pair à pair.

**Contributions et analyse du document** Après une brève introduction situant le contexte et les enjeux de son travail, S. Martin dresse, dans le chapitre 2, un panorama des éditeurs collaboratifs et recense une série de propriétés spécifiques de ces outils. En particulier, il

précise clairement le sens de la propriété qui va être le fil conducteur de son travail : *la convergence*.

Le chapitre 3 est consacré à l'étude détaillée de l'état de l'art en matière d'édition pair à pair. Le candidat mentionne que selon la nature des documents (texte simple, documents structurés, arborescents ou au format XML) les techniques vont être différentes et plus ou moins adéquates. Ensuite, deux grandes classes d'approches adaptées pour la résolution du problème sont décrites : l'approche par transformées opérationnelles et l'approche CRDT (Commutative Replicated Data Type). La première approche consiste à doter chaque participant d'un *programme d'intégration* ce programme reçoit les opérations effectuées par les autres participants et tente de les intégrer (par réécriture/recombinaison) de sorte à prendre en compte ce qui a été fait par les autres collaborateurs. Si le programme intégrateur satisfait deux propriétés (TP1 et TP2) alors l'éditeur collaboratif assurera la *convergence*. La seconde classe d'approche exploite le fait que certaines opérations d'édition commutent. Ainsi l'ordre selon lequel elles sont exécutées, d'un point de vue sémantique, est indifférent. Ces opérations sont donc parallélisables.

Le chapitre 4 du manuscrit s'intéresse aux transformées opérationnelles. Pour différentes structures de données de plus en plus complexes (arbres d'arité non bornée étiquetés en leurs arêtes, arbres avec identifiants uniques et enfin graphes fonctionnels), et des jeux d'opérations variés dans un ensemble donné (*Del1*, *Del2* destruction, *Add* ajouts, *Move* déplacement d'un sous-arbre, *Nop* opération laissant l'arbre inchangé), S. Martin va établir des résultats à la fois négatifs et des résultats d'obtention d'intégrateurs qui satisfont les propriétés TP1 et TP2.

Plus précisément, une première contribution (section 4.2), est de proposer une structure de données de document semi structurés similaire à celle proposée par Pierce et al. [PSG03] avec des opérations d'édition de base permettant la définition de transformées opérationnelles. Il est ainsi montré que cette structure munie du jeu d'opérations {*Nop*, *Add*, *Del2*} ne permet pas d'obtenir un éditeur collaboratif, quelle que soit la transformée opérationnelle utilisée (aucune transformée ne peut satisfaire TP1 et TP2 dans ce cas). En revanche, pour la même structure munie des opérations {*Add*, *Del1*, *Nop*} l'algorithme de Ressel s'applique est il devient possible de définir une transformée opérationnelle.

Une structure de données un peu plus sophistiquée qui utilise des identifiants uniques pour l'étiquetage des arêtes est ensuite proposée. Ces identifiants uniques permettent des accès non ambigus dans le document. La gestion de tels identifiants n'est pas coûteuse et n'induit pas de surcoût pour la manipulation du document. Pour cette structure de données S. Martin propose différents jeux d'opérations et des transformées opérationnelles qui garantissent les propriétés TP1 et TP2 requises pour la convergence. Il montre également un résultat négatif très instructif : l'incompatibilité des opérations *Del1* et *Del2*.

Le traitement spécifique de l'opération *Move* est entrepris en section 4.6 et d'autres résultats d'incompatibilité sont obtenus entre *Move* et *Del2*.

Enfin, S. Martin propose une généralisation des arbres aux graphes fonctionnels (graphes orientés de degré sortant maximal valant 1) qui permet également de prendre en compte l'opération de déplacement (*Move*) d'un sous-arbre avec une méthode de transformée opérationnelle.

J'ai apprécié le fait que le candidat ait fait appel au démonstrateur *Vote*, dédié à ces propriétés, pour établir, dans de nombreux cas, les preuves de ses théorèmes.

Au chapitre 5, S. Martin propose la conception d'un algorithme générique de type CRDT, *FCedit*, qui assure la propriété de convergence grâce à une notion de *dépendance sémantique*. Cet algorithme ne se base pas sur l'historique (ce qui est coûteux), il passe donc à l'échelle,

et a été validé par un prototype.

Au chapitre 6, une structure de données pour les documents semi-structurés et un jeu d'opérations satisfaisant la propriété d'indépendance requise par l'algorithme pour assurer la convergence est présentée. La modélisation (et la structure de donnée en question) permet de prendre en compte les documents au format XML. Les résultats obtenus au chapitre précédent sont étendus au cas de cette structure de données. En outre, la prise en compte d'opérations d'annulation est assurée ce qui constitue une réelle avancée dans ce cadre. Enfin il ressort de ce chapitre qu'il est possible d'obtenir un algorithme d'édition collaborative, efficace, pour les documents au format XML : cela aura un impact pratique certain à l'avenir.

Dans le chapitre 7, une nouvelle contribution consiste à prendre en compte le typage dans le processus d'édition, en adaptant l'algorithme générique de manière à garantir que la convergence soit établie et ce en enrichissant le langage d'expression des types (DTD, Schemas et extended general tree grammars). L'idée est que lors du processus d'édition, le type du document n'est pas préservé. Le candidat propose une opération de synchronisation ainsi que des algorithmes de réparation (du type) pour garantir que les documents obtenus soient bien typés. Il est montré que la solution est applicable pour de nombreux schémas de type pourvu qu'il existe un algorithme de réparation pour les types considérés utilisant les opérations d'édition classiques (cet algorithme calcule les opérations d'édition à effectuer pour que le document soit bien typé).

Le travail présenté au chapitre 8, complète les résultats de l'approche CRDT car prenant en compte l'opération de déplacement (*Move*) d'un sous-arbre en présence des opérations classiques. Ceci est rendu possible en passant par une structure de données de type graphe. Les documents semi-structurés correspondent alors à un cas particulier de tels graphes. La généralisation des résultats précédents est de nouveau donnée. Ceci n'était pas pris en compte dans les travaux antérieurs.

Le chapitre 9, dernière contribution de cette thèse est la réalisation d'un prototype qui implémente les solutions décrites aux chapitres précédents, en vue de les valider.

### **Conclusion**

Pour conclure, sur le fond, le manuscrit présenté par S. Martin est scientifiquement de très bonne teneur. Sur la forme, il ressort que les qualités littéraires du candidat ne sont pas son atout majeur. Cependant, les résultats sont solides et apportent une contribution indéniable à l'état de l'art du domaine. J'ai particulièrement apprécié le bon équilibre entre formalisation et implantation au sein d'un outil. Les travaux présentés dans cette thèse ont dorénavant et déjà donné lieu et donneront lieu à des publications de niveau international, j'en suis convaincue. Pour toutes ces raisons je recommande vivement que le titre de docteur en informatique de l'Université Aix-Marseille soit décerné à Monsieur Stéphane Martin.





## A.6 Rapports de présoutenance de Pascal Molli

---



### **Rapport sur la thèse de Stéphane Martin “édition collaborative des documents semi-structurés”**

La thèse de Stéphane Martin intitulée “édition collaborative des documents semi-structurés” s’inscrit dans le contexte des systèmes collaboratifs distribués. Ce domaine de recherche s’est développé principalement au sein des communautés CSCW/CHI (Computer Supported Cooperative Work/Computer Human Interaction), Systèmes distribués (PODC/TPDS), Bases de données (Sigmod). La thèse s’attaque aux problèmes de gestion de la cohérence des données réparties dans un environnement distribué décentralisé sur des données semi-structurées, en l’occurrence XML. Ce sujet est d’une grande actualité poussé par le développement sans précédent du web social.

Le manuscrit est divisé en 10 chapitres concis:

1. Le chapitre 1 définit le contexte, les enjeux de la thèse ainsi que les différentes contributions présentées dans le manuscrit.
2. Le chapitre 2 présente les différents concepts utilisés dans le domaine de l’édition collaborative distribuée. Ce chapitre cerne bien l’ensemble des concepts centraux des systèmes collaboratifs distribués ainsi que les outils les mettant en oeuvre. Certains concepts sont particulièrement bien présentés comme celui des dépendances sémantiques, d’autres concepts comme la préservation des intentions ou l’annulation de groupe auraient mérités la même attention.
3. Le chapitre 3 présente les différents types de données utilisés dans la thèse ainsi que les algorithmes existant permettant l’édition collaborative distribuée de ces données. Ce chapitre introduit les principes de l’approche des transformées opérationnelles (OT) et des types commutatifs répliqués (CRDT). Ce chapitre montre bien certaines limites des approches existantes : difficulté des preuves de convergence, difficulté de supporter les architectures P2P, travaux surtout orientés sur le support de structures linéaires. Ces critiques sont fondées et motivent les contributions des 5 chapitres suivant. L’algorithme COT (CSCW06,TPDS09) aurait du être référencé dans cette partie.

4. Le chapitre 4 présente une première contribution dans le cadre de l'approche des transformées opérationnelles avec la présentation de jeux de transformées pour des structures arborescentes. Les types abstraits considérés sont bien formalisés et les preuves des propriétés TP1 et TP2 sont crédibles. Les preuves d'impossibilité d'existence de fonction de transformations sont une contribution originale de ce chapitre. La manière de prouver ces impossibilités est très intéressante et très innovante dans le domaine.
5. Le chapitre 5 présente l'algorithme FCedit. Cet algorithme s'inscrit résolument dans l'approche des types commutatifs répliqués. L'originalité de l'algorithme repose sur une réception "causale sémantique" qui rend l'algorithme générique. Une opération est délivrée si ses dépendances sémantiques sont vérifiées. Sous l'hypothèse que les opérations du type abstrait sont indépendantes, FCedit garantit la convergence. Ce chapitre est bien formalisé, les définitions sont claires, les preuves sont limpides. FCedit est ensuite instancié sur les arbres étiquetés. Cela permet de bien comprendre comment l'algorithme fonctionne. Les limites de l'approche sont expliquées clairement en fin de chapitre. FCedit est intéressant dans sa formalisation. Cependant, d'autres approches sont aussi fondées sur des dépendances sémantiques notamment WOOT et les travaux de Ladin en 92. Il aurait été intéressant de positionner FCedit avec ces approches.
6. Le chapitre 6 est une instanciation de l'algorithme FCedit pour les arbres XML ordonnés. L'idée est d'encoder les relations d'ordre dans les étiquettes des arêtes. Cette approche est basée sur l'idée de l'algorithme Logoot-Undo. Ce chapitre est bien formalisé, les preuves sont claires. Ce chapitre peut être vu comme une combinaison originale et intéressante des algorithmes FCedit et Logoot-Undo.
7. Le chapitre 7 aborde les problèmes de respect des grammaires dans l'édition collaborative. En effet, l'édition concurrente de documents XML peut facilement entraîner des violations de DTD. L'approche proposée repose sur des algorithmes de réparations existant. Chaque site maintient une vue globale et une vue locale d'un document. La vue globale est convergente mais peut être incorrecte, les vues locales sont produites par les algorithmes de réparations et doivent converger vers une version correcte du document. Ce chapitre est assez confus dans ses explications ainsi que dans son positionnement par rapport à l'état de l'art. Cependant, l'idée générale est intéressante et l'algorithme proposé est correct.
8. Le chapitre 8 propose un CRDT pour les graphes. La formalisation est claire et les preuves de correction convaincantes. Le type abstrait est ensuite étendu avec l'opération "move". Ce chapitre est intéressant dans

- la mesure ou il démontre comment il est possible de produire de nouveaux CRDT et prouver leurs corrections. L'approche "systématique" utilisée dans ce chapitre peut être utilisée pour produire d'autres CRDT.
9. Le chapitre 9 présente une implantation de l'algorithme FCedit ainsi que quelques résultats expérimentaux. Ce chapitre est très court. Il est dommage que les expérimentations n'inclue pas d'éléments de comparaison. Il aurait été intéressant de comparer l'impact sur les performances entre une approche basée sur une réception causale classique et une réception causale sémantique. Cependant, les premiers résultats expérimentaux sont encourageants.
  10. Le chapitre 10 conclue la thèse et présente quelques perspectives.

La thèse dans son ensemble présente des contributions algorithmiques nombreuses et variées dans le domaine des systèmes collaboratifs distribués. Les structures de données utilisées ainsi que les algorithmes proposés sont clairement formalisés et prouvés. Certains résultats ont été validés par des publications internationales (OTM 2010, CDVE 2010). La thèse ouvre également des perspectives intéressantes en montrant comment écrire de nouveaux types abstraits garantissant la convergence.

Pour les raisons exprimées ci-dessus, j'estime que le travail présenté dans ce manuscrit permet la soutenance de la thèse intitulée "édition collaborative des documents semi-structurés" présentée par Stéphane Martin pour l'obtention du titre de Docteur de l'Université d'Aix-Marseille.

**Pascal Molli**  
**Professeur à l'Université de Nantes,**  
**à Nantes, le 18 Juillet 2011**

