

CURRICULUM-VITAE

	Hubert LARCHEVÊQUE
Né le	23 Juillet 1985 à Ivry sur Seine (94)
Nationalité	Française
Position	POST-DOCTORANT
Adresse	Campus Universitaire 351, cours de la libération 33405 Talence cedex France
Courriel	Tel. : (+33) 5 4000 3883 Hubert.Larcheveque@labri.fr
Page	http://www.labri.fr/perso/larcheve

Diplômes et Expérience Professionnelle

- 2011-2012 **Post-doctorat** réalisé à Nantes (EPC Ascola) et Bordeaux (EPC Cepage), financé par l'ANR USS-Simgrid.
- 2010-2011 **Demi-ATER** à l'ENSEIRB.
- 2007-2010 **Doctorat en Informatique** au sein de l'équipe **Cepage** (INRIA Bordeaux - Sud-Ouest), titulaire d'une bourse BDI CNRS-Région Aquitaine.
- “Agrégation de ressources avec contrainte de distance : applications aux plateformes de grande échelle”.
- Directeurs : Olivier Beaumont, Nicolas Bonichon, Philippe Duchon.
 - Rapporteurs : Christian Laforest, Christian Lavault.
 - Jury : Laurent Viennot (président), Christian Laforest, Christian Lavault, Anne Benoit.
- 2007-2010 **Moniteur** de l'Université de Bordeaux : 245h de cours (6h de cours, 126h de TDs, 113h de TPs). Public : de L1 à dernière année d'école d'ingénieur.
- Fév-Juin 2007 **Stage de Master Recherche** à l'Université de Bordeaux, avec Olivier Beaumont, Nicolas Bonichon et Philippe Duchon. “Bin covering distribué avec et sans contrainte de distance pour les plateformes à grande échelle”.
- 2006-2007 **Master Recherche** à l'Université de Bordeaux : “Modèles pour l'Algorithmique Distribuée”.
- 2004-2007 **Diplôme d'ingénieur**(Informatique) : ENSEIRB, option “Parallélisme, Régulation et Calcul Distribué”.

1 Centres d'intérêts en recherche

Mots clés : Bin covering, bin packing, algorithmes d'approximation, augmentation de ressources, algorithmes distribués, structure de données probabiliste, réseaux de grande échelle, plongement d'Internet, placement de serveurs/répliqués, allocation de ressources, Cloud Computing.

J'ai effectué mes études d'ingénieur parallèlement à mon master recherche à l'université de Bordeaux, ainsi que mon stage de Master qui a abouti sur une thèse. Pendant ma thèse, j'ai effectué un court séjour à l'Université de Wrocław, en Pologne, pour y travailler aux côtés de Mirosław Korzeniowski. Après ma thèse, j'ai occupé un poste d'ATER à l'Institut Polytechnique de Bordeaux avant de commencer un post-doctorat à Nantes, au sein de l'équipe Ascola, sur un nouveau sujet ayant trait à des problèmes d'allocations de ressources dans le domaine du Cloud Computing.

1.1 *Bin Covering* avec Contrainte de Distance

Bin Covering est un problème NP-difficile classique : étant donné un ensemble d'éléments pondérés, l'objectif est de construire un nombre maximal de groupes d'éléments, chacun ayant un poids total supérieur à une

certaine valeur fixée. Nous avons ajouté une contrainte additionnelle à ce problème, en plaçant les éléments dans un espace métrique. Deux éléments placés dans le même groupe doivent alors se trouver à une distance inférieure à une certaine valeur d_{max} . Pour étudier ce problème nous avons eu recours à de l'augmentation de ressources : nous avons autorisé nos algorithmes à construire des groupes de diamètre au plus βd_{max} en comparant le nombre de groupes ainsi créés au nombre de groupes d'une solution optimale contrainte à construire des groupes de diamètre d_{max} . Cette augmentation de ressources est nécessaire puisque nous avons prouvé que, pour tout $\epsilon > 0$, il est NP-difficile d'élaborer un algorithme s'exécutant en temps polynomial et assurant un facteur d'approximation constant pour ce problème en dépit de l'utilisation d'une augmentation de ressources de $(2 - \epsilon)$. Nous avons par ailleurs démontré que quelle que soit l'augmentation de ressources utilisée, il est NP-difficile d'élaborer un algorithme s'exécutant en temps polynomial et assurant un facteur d'approximation supérieur à $1/2$ comparé à une solution optimale contrainte à d_{max} . Nous avons présenté un algorithme s'exécutant en temps polynomial assurant un facteur d'approximation de $2/5$, et une version distribuée de cet algorithme assurant une approximation de $1/3$, tout deux utilisant une augmentation de ressources de 4. Nous avons également étudié ce problème dans plusieurs espaces métriques spécifiques, utilisés par divers outils existants pour modéliser les réseaux de grande échelle visés par nos applications. Nous avons en particulier obtenu des résultats intéressants dans un plan (muni d'une norme quelconque) et dans un arbre.

1.2 Sur la hauteur des skip-graphs aléatoires

La structure de données que nous utilisons pour ordonner nos éléments dans la version distribuée de l'algorithme pour le *Bin Covering* avec Contrainte de Distance est un skip-graph, une variante de la structure de donnée appelée skip-list, destinée à offrir de meilleures performances dans un environnement distribué. Pour s'assurer des bonnes performances de notre algorithme distribué, nous avons dû étudier la hauteur des skip-graphs. Cette hauteur correspond à la longueur maximale d'un chemin de recherche depuis n'importe quel élément, vers n'importe quel autre élément dans un skip-graph. Nous avons prouvé qu'elle est d'ordre $\log n$ dans un skip-graph contenant n éléments.

1.3 *Bin Packing* avec Contrainte de Distance

Bin Packing est un autre problème NP-difficile classique : étant donné un ensemble d'éléments pondérés, l'objectif est de minimiser le nombre de groupes ne dépassant pas un certain poids nécessaires à ce que chaque élément appartienne à un groupe. Nous avons ajouté la contrainte additionnelle équivalente plaçant les éléments dans un espace métrique. Deux éléments d'un même groupe doivent alors se trouver à une distance inférieure à d_{max} l'un de l'autre. Pour étudier ce problème nous avons eu recours à l'augmentation de ressources de la même façon que dans le cas du *Bin Covering*. Nous avons prouvé des résultats d'inapproximabilité équivalents, à savoir que pour tout $\epsilon > 0$, il est NP-difficile d'élaborer un algorithme s'exécutant en temps polynomial et assurant un facteur d'approximation constant pour ce problème en dépit de l'utilisation d'une augmentation de ressources de $(2 - \epsilon)$, et quelle que soit l'augmentation de ressources utilisée, il est NP-difficile d'élaborer un algorithme s'exécutant en temps polynomial et assurant un facteur d'approximation inférieur à $3/2$ comparé à une solution optimale contrainte à d_{max} . Nous avons présenté un algorithme s'exécutant en temps polynomial assurant un facteur d'approximation de $7/3$ utilisant une augmentation de ressources de 2. Nous avons également étudié ce problème dans des espaces métriques particuliers comme le plan euclidien et l'arbre binaire.

Enfinement nous avons présenté un résultat de correspondance liant le *Bin Packing* avec Contrainte de Distance au problème de K -Centre avec Capacités, une autre variante d'un problème NP-difficile classique. En particulier, à notre connaissance, nos résultats pour le *Bin Packing* avec Contrainte de Distance impliquent les premiers résultats connus concernant la variante non uniforme du K -Centre avec Capacités.

1.4 Comparaison entre différents outils de plongement d'Internet

Il est connu que les réseaux de grande échelle, comme Internet, ont une structure bien particulière. De nombreux outils ont pour objectif de plonger ce type de grands réseaux dans des espaces métriques bien décrits, dans lesquels il peut être plus simple de travailler. C'est le cas en particulier de Sequoia, qui plonge Internet dans un arbre dont les feuilles sont les noeuds du réseau, et de Vivaldi, qui plonge un réseau dans une généralisation du plan euclidien.

Nous nous sommes donc intéressés à l'étude de nos problèmes dans ce type d'espaces métriques, plus simples et semblant correspondre plus exactement aux réseaux qui nous intéressent. En pratique, l'évaluation de la modélisation du réseau considéré doit tenir compte de l'application pratique visée. Ainsi nous avons étudié comment la simplicité de l'espace métrique de plongement joue un rôle important dans l'élaboration d'algorithmes, aboutissant à des résultats d'approximation meilleurs, car exploitant les spécificités de ces espaces métriques plus simples induits par chaque outil de plongement. Via des simulations nous avons été en mesure d'identifier quels couples algorithmes/outils de plongement étaient les plus efficaces pour chacune des applications étudiées.

1.5 Placement de serveurs avec contraintes de distance dans des arbres

Les problèmes de placements de serveurs à capacité finie dans des réseaux sont des problèmes bien connus et bien étudiés de manière générale. Généralement, un ensemble de clients se trouve placé dans un réseau, et chaque client a une certaine quantité de requêtes qu'un serveur doit traiter. L'objectif est de placer un minimum de serveurs de capacité fixe de façon à ce que toutes les requêtes soient satisfaites, en respectant certaines contraintes de distance : dans notre cas, un client ne doit pas être à plus d'une certaine distance des serveurs traitant ses requêtes. Nous nous sommes intéressés aux cas particuliers où l'ensemble de clients est l'ensemble des feuilles d'un arbre, dont les noeuds internes peuvent héberger des serveurs. Nous avons étudié plusieurs variantes de ce problème, l'une dans laquelle toutes les requêtes d'un client doivent être traitées par un même serveur, tandis que dans l'autre variante, plusieurs serveurs peuvent être utilisés pour traiter les requêtes d'un même client.

Nous avons prouvé que ce problème, dans sa première variante, est NP-difficile, et avons proposé des algorithmes d'approximation pour le résoudre, notamment un algorithme s'exécutant en temps polynomial assurant un facteur d'approximation de 2 lorsque la contrainte de distance n'est pas considérée. Dans sa seconde variante, nous avons prouvé que, dans le cas particulier des arbres binaires, il existe un algorithme s'exécutant en temps polynomial permettant de résoudre ce problème de manière optimale.

1.6 Allocation de ressources pour le Cloud Computing

Dans le domaine du Cloud Computing, de nombreux problèmes d'optimisation de la répartition des ressources se posent. Dans ce domaine, des utilisateurs louent des ressources, identifiées sous le nom de "machines virtuelles", possédant chacune des caractéristiques spécifiques aux desiderata des utilisateurs en terme de puissance de calcul, d'espace de stockage, etc ... Ces machines virtuelles sont louées à des propriétaires qui disposent d'un ensemble de machines physiques, aux capacités définies. L'objectif pour un propriétaire de "machines physiques" est d'optimiser l'utilisation de son parc de machines, pour pouvoir gérer le plus d'utilisateurs possibles. Il lui faut donc, dans un premier temps, répartir convenablement les machines virtuelles demandées par ses clients sur ses machines physiques, de façon qu'à tout moment chaque utilisateur puisse avoir effectivement à sa disposition une machine virtuelle de la capacité qu'il a choisi.

Dans un deuxième temps, le propriétaire peut chercher à effectuer ce qu'on nomme de la consolidation : ce procédé exploite le fait qu'un utilisateur ayant loué une machine virtuelle n'utilisera pas les capacités de celle-ci au maximum de leur potentiel tout le temps, mais que bien plus sûrement il a choisi une telle machine virtuelle de façon à être sûr qu'elle sera capable de supporter la charge de travail requise lors d'un éventuel pic de charge lié à son activité (le cas des serveurs webs en est un bon exemple : ils ont souvent un pic journalier et leur charge de travail le reste de la journée est relativement faible comparée à ce pic). Le propriétaire peut alors décider de n'allouer à chaque machine virtuelle que ce qu'elle utilise réellement, en faisant en sorte que dès qu'elle aura besoin de plus grandes capacités, dans la limite accordée au client, elle puisse y avoir accès, dans l'idéal immédiatement. L'utilisation de ce procédé nécessite l'utilisation d'algorithmes d'allocation de ressources à la fois robustes et réactifs, susceptibles de réagir de manière rapide et efficace à des changements impromptus de la demande de ressources des machines virtuelles hébergées.

Nous travaillons actuellement, au sein d'une collaboration liant une équipe nantaise (Ascola) et une équipe bordelaise (Cepage), à l'élaboration d'algorithmes réactifs et robustes qui permettraient l'allocation de ressources multi-critères à destination des plateformes de Cloud Computing, destinés à être intégrés à des outils déjà existants.

2 Enseignement

Tableau résumant mes activités d'enseignement :

Intitulé	Volume Horaire	TD/TP/Cours	Niveau
Analyse syntaxique	35h	TDs et TP	L3
Environnement de Développement	18h	TPs	L2
Utilisation des Réseaux	35h	TPs	L3
Algorithmique et programmation	27h	TDs	1ère année MATMECA
Certificat "Informatique et Internet"	44h	TPs	L1
Initiation à l'Informatique	32h	TDs et TP	L1
Compilation	16h	TDs	2ème année ENSEIRB
Probabilités et Statistiques	16h	TDs	1ère année ENSEIRB
XML et Technologies Web	16h	TDs et TP	2ème année ENSEIRB
Outils de modélisation d'Internet	6h	Cours	3ème année ENSEIRB

J'ai été pendant trois ans (2007-2010) moniteur à l'Université Bordeaux 1. Durant ces trois années j'ai encadré les travaux pratiques et dirigés dans les matières suivantes :

- Analyse Syntaxique, 35h, Licence 3 : analyse lexicale et syntaxique, compilateur, ...
- Environnement de Développement, 18h, Licence 2 : utilisation du compilateur, d'outils de debug, d'emacs avancée, d'outils de versionnage, ...
- Utilisation des Réseaux : 35h, Licence 3 : protocoles d'internet (ICMP, TCP, IP, UDP), programmation de sockets, outils pour la programmation réseau (valgrind, wireshark, ...), ...
- Algorithmique et Programmation, 27h, première année de l'école d'ingénieur MATMECA : introduction à l'algorithmique (algorithmes de tri ...), aux structures de données de base (arbre, tas, ...), implémentation en Fortran : responsabilité de la création de l'intégralité des sujets de TDs, partagée avec les autres doctorants encadrant les mêmes TDs.
- Certificat "Informatique et Internet", 44h, Licence 1 : introduction à l'utilisation d'une espace virtuel de travail, internet, manipulation de différents logiciels de bureautique (powerpoint, excel, word), ...
- Initiation à l'Informatique, 32h, Licence 1 : implémentation d'algorithmes basiques sur les graphes (plus courts chemins, connexité, coloration, ...) en Python.

En tant qu'ATER à l'Institut Polytechnique de Bordeaux, j'ai également encadré les matières suivantes :

- Compilation, 16h, deuxième année de l'ENSEIRB : responsabilité de la création des derniers sujets de TDs, automates des items, P-code.
- Probabilités et Statistiques, 16h, première année de l'ENSEIRB : lois de probabilités discrètes, à densité, processus markovien, ...
- XML et Technologies Web, 16h, deuxième année de l'ENSEIRB : langage XML, librairie SAX, DOM, XPath, XSL, ...

J'ai également, deux années de suite (2010 et 2011), donné un cours concernant les **outils de modélisation d'Internet** dont la création était entièrement à ma charge, destiné aux élèves de la filière Parallélisme, Régulation et Calculs Distribués en troisième année à l'ENSEIRB.

3 Liste de Publications

Références

3.1 Articles publiés dans des conférences internationales avec comités de lecture

- [1] O. Beaumont, N. Bonichon, P. Duchon, L. Eyraud-Dubois, and H. Larchevêque. A Distributed Algorithm for Resource Clustering in Large Scale Platforms. In *Principles of Distributed Systems (OPODIS)*, pages 564–567. Springer, 2008.

- [2] O. Beaumont, N. Bonichon, P. Duchon, and H. Larchevêque. Distributed approximation algorithm for resource clustering. In *Structural Information and Communication Complexity (SIROCCO)*, pages 61–73. Springer, 2008.
- [3] P. Duchon, and H. Larchevêque. On the search path length of random binary skip graphs. In *Analytic Algorithmics and Combinatorics (ANALCO10)*, SIAM, 2010, ISBN : 978-0-898719-33-8, pages 1–8.
- [4] O. Beaumont, N. Bonichon, P. Duchon and H. Larchevêque. Use of Internet Embedding Tools for Heterogeneous Resources Aggregation. In *Heterogeneity in Computing Workshop (HCW2011)*, IEEE, 2011.
- [5] O. Beaumont, N. Bonichon and H. Larchevêque. Modeling and Practical Evaluation of a Service Location Problem in Large Scale Networks. In *International Conference on Parallel Processing (ICPP2011)*,
- [6] A. Benoit, H. Larchevêque and P. Renaud-goud. Optimal algorithms and approximation algorithms for replica placement with distance constraints in tree networks. To appear in *International Parallel and Distributed Processing Symposium (IPDPS2012)*, IEEE, 2012.

3.2 Articles publiés dans des conférences nationales avec comités de lecture

- [7] O. Beaumont, N. Bonichon, P. Duchon and H. Larchevêque. Bin Packing avec Contrainte de Distance. ROADEF 2011.
- [8] O. Beaumont, N. Bonichon, P. Duchon and H. Larchevêque. Utilisation d’outils de plongement d’Internet pour l’agrégation de ressources hétérogènes. In *13es Rencontres Francophones sur les Aspects Algorithmiques de Télécommunications (AlgoTel2011)*,

3.3 Rapports de Recherche

- [9] O. Beaumont, N. Bonichon, P. Duchon, L. Eyraud-Dubois, and H. Larchevêque. A Distributed Algorithm for Resource Clustering in Large Scale Platforms. Research report N°6883, INRIA Bordeaux Sud-Ouest, France, October 2008.

4 Responsabilités Administratives

- Relecteur de plusieurs articles pour différentes conférences internationales (IPDPS 2010 et 2011, ...).
- 2009-2010 : **Membre Actif** du Comité d’Organisation de la conférence MajecSTIC’10, tenue à Bordeaux en Octobre 2010.
- 2008-2009 : **Membre du Comité de Relecture** de la conférence MajecSTIC’09.
- 2007-2008 : **Président** de l’association AFoDIB, association des doctorants du LaBRI, Laboratoire Bordelais de Recherche en Informatique.

5 Loisirs

- 2007-2010 : **Musique** : saxophone dans les fanfares bordelaises “La Grasse Bande” et “Los Teoporos”.
- 2008-2010 : **Président et Vice Président** de l’association du groupe musical “La Grasse Bande”.
- 2008-2009 : **Trésorier** de l’association “Union Libre 2009”, qui a organisé un festival de musique à Bordeaux durant l’été 2009.
- 2005-2006 : **Trésorier** du Bureau des Sports de l’ENSEIRB, qui a géré la commission Sports du TIE 2006 (Tournoi sportif réunissant des écoles d’ingénieur françaises).

6 Langues Etrangères

- Anglais : courant.
- Espagnol : bonnes notions, écrit, lu.
- Italien : notions basiques.