

Inauguration du PCRI

Mercredi 9 novembre 2011

Quelques éléments d'informations organisationnelles et scientifiques

Le bâtiment PCRI (Pôle Commun de Recherche en Informatique), dont le maître d'ouvrage a été l'Université Paris-Sud, est partagée entre l'Université Paris-Sud et Inria et héberge deux entités : la quasi-totalité des équipes du LRI (Laboratoire de Recherche en Informatique, unité mixte de recherche de l'Université Paris-Sud et du CNRS, rattaché à l'INS2I) et plusieurs équipes-projets du centre Inria Saclay – Île-de-France, beaucoup de ces équipes étant communes. Y sont ainsi regroupés des enseignants-chercheurs et BIATSS de Paris-Sud, des chercheurs et ITA CNRS, des chercheurs et assistant(e)s de projets Inria, ainsi que les secrétariats pédagogiques du master 2 recherche en Informatique et de l'École Doctorale en Informatique de Paris-Sud, soit de l'ordre de 300 personnes (dont plus de 100 permanents).

Le LRI, qui a plus de 30 ans d'existence, était historiquement situé au bâtiment 490 sur le Campus d'Orsay-Vallée de l'Université Paris-Sud. Au début des années 2000, le laboratoire a commencé à nouer une étroite collaboration avec le Centre Inria Futurs, qui a donné naissance à Inria Saclay – Île-de-France en janvier 2008. Des équipes communes LRI-Inria ont été formées : certaines, ainsi qu'une équipe-projet propre Inria, abritées au 490 ; d'autres, ainsi qu'une équipe propre LRI, hébergées dans les locaux Inria du Parc Club sur le plateau. Cette dispersion géographique n'était pas favorable à la cohérence scientifique de l'ensemble. Depuis juin 2011, toutes ces équipes ont eu la satisfaction d'être regroupées au sein du PCRI, à l'exception d'une équipe commune qui intégrera au premier trimestre 2012 le nouveau bâtiment de Digiteo Moulon, connexe au PCRI, achevant ainsi la réunion de tous en un même lieu.

Inria ayant signé des accords-cadres d'une part avec Paris-Sud (puis, plus généralement avec la CPU), d'autre part avec le CNRS, ceux-ci vont être déclinés et spécifiés prochainement au niveau LRI - Inria Saclay pour régir le statut et l'organisation des équipes communes. Ce ne sera pour la plupart des points qu'une institutionnalisation des bonnes pratiques de recherche et vie communes existant depuis longtemps entre les deux unités.

Les thèmes de recherche des équipes du PCRI couvrent un large spectre de l'informatique et incluent à la fois des aspects fondamentaux et des aspects appliqués : algorithmique, complexité, théorie des graphes, optimisation combinatoire, génie logiciel, test et preuve de programmes, parallélisme, clusters et grilles de calcul, réseaux, bases de données, intelligence artificielle et systèmes d'inférence, fouille de données et apprentissage statistique, interaction homme-machine, visualisation de données, bioinformatique, etc. Cette diversité est l'une des forces du PCRI car elle favorise les recherches aux frontières, là où le potentiel d'innovation est le plus grand.

Plus précisément, aujourd'hui, le PCRI héberge les équipes administrative et technique du LRI (Université Paris-Sud/CNRS) et onze de ses douze équipes de recherche, dont sept sont communes avec des équipes-projets Inria, ainsi enfin qu'une équipe-projet Inria propre.

Les huit équipes communes (totalement ou partiellement) au LRI (Université Paris-Sud/CNRS) et à Inria sont :

- **Proval : Preuves de programme**

L'objectif de ces recherches est de proposer des méthodes et des outils pouvant s'intégrer dans le cycle de développement logiciel et permettant de produire du code correct vis-à-vis de comportements attendus. Il s'agit de développer une plateforme générique de vérification de programmes, Why, qui produit des obligations de preuve à partir de programmes annotés de spécifications formelles et les transmet ensuite à des outils de preuve, interactifs ou automatiques. Des instances pour les langages C (Caduceus) et Java (Krakatoa) sont bâties sur cette infrastructure générique.

- **Parall (nom de l'équipe LRI) / Grand Large (nom de l'équipe Inria) : Parallélisme**

Les thèmes de recherche abordés sont d'une part l'étude et la conception d'algorithmes répartis ainsi que la modélisation et la vérification des systèmes répartis, d'autre part la construction et l'expérimentation d'un environnement de calcul global pair à pair fonctionnant sur Internet, Intranet et LAN, le traitement de la volatilité et la fiabilité du point de vue fondamental (adéquation de l'auto-stabilisation, protocoles niveau réseau, vérification de protocoles) et la programmation et méthodes numériques pour le calcul à grande échelle (HPC).

- **Archi (nom de l'équipe LRI) / Alchemy (nom de l'équipe Inria) : Architectures parallèles**

Cette équipe étudie les architectures, langages et compilateurs pour des calculateurs embarqués non spécifiques de haute performance. Les techniques utilisées sont les optimisations de programme à base de compilation dynamique ou itérative, l'approche combinée langage / architecture avec passage d'une sémantique plus riche du langage vers l'architecture et le compilateur, la simulation.

- **BD + IASI (noms des équipes LRI) / Leo (nom de l'équipe INRIA) : Bases de données + Intelligence artificielle et systèmes d'inférence**

La plus grande partie de l'équipe BD, dont les thèmes de recherche sont les bases de données multidimensionnelles ou temporelles, le stockage et l'optimisation, l'intégration d'informations, les bibliothèques numériques, le langage XML, a rejoint récemment, au sein de l'équipe Inria Leo, l'équipe IASI qui travaille sur le problème de l'intégration d'informations distribuées et hétérogènes, devenu un enjeu crucial pour l'exploitation de l'immense gisement d'informations rendu accessible par le Web. Ensemble, ces deux équipes conduisent une recherche coordonnée sur les techniques formelles pour modéliser et organiser les données, les connaissances et le raisonnement, sur l'exploitation et l'enrichissement de la sémantique des données dans l'objectif de l'interopérabilité et sur la conception d'algorithmes et le déploiement de systèmes qui répondent efficacement aux besoins complexes et croissants de gestion de données.

- **In Situ : Interaction située**

Ce projet de recherche vise à développer de nouvelles techniques d'interaction, de nouveaux outils pour mettre en œuvre ces techniques, et de nouvelles méthodologies pour maîtriser le processus de conception. Il s'agit de créer des interfaces situées, c'est-à-dire adaptées (ou adaptables) à leurs contextes d'usage et exploitant au mieux la complémentarité entre systèmes informatiques et utilisateurs humains. A terme, il s'agit de réaliser une nouvelle génération d'environnements interactifs destinés à supplanter les interfaces graphiques actuelles.

- **BioInfo (nom de l'équipe LRI) / AMIB (nom de l'équipe Inria) : Bioinformatique**

Cette équipe de recherche pluridisciplinaire rassemble des compétences complémentaires en informatique, telles que algorithmique, combinatoire, fouille de données, intégration de données, modélisation et simulation, et travaille en étroite collaboration avec des biologistes, notamment avec l'Institut de Génétique et Microbiologie (Université Paris-Sud/CNRS). Elle s'intéresse principalement à

déterminer l'organisation de l'information dans les séquences d'ADN et à identifier le rôle joué par les produits des gènes que sont les ARNs et les protéines, en même temps que leur dynamique dans les processus cellulaires. Toutes les fonctions des protéines et presque toutes celles des ARN étant reliées à leur structure moléculaire 3D, la prédiction de structures et l'analyse des séquences sont au cœur de l'annotation des gènes. Les axes de recherche de l'équipe sont ainsi les approches algorithmiques des structures biologiques, la fouille de données biologiques, l'intégration de données massives pour l'annotation génomique et la biologie des systèmes.

- **A&O (nom de l'équipe LRI) / TAO (nom de l'équipe Inria) : Apprentissage et optimisation**

La fertilisation croisée de l'apprentissage automatique, de la découverte de connaissances et du calcul évolutionnaire est la motivation première de cette équipe. Elle mène ses recherches selon six axes fortement couplés : calcul autonome (spécialisé en particulier au contexte des grilles à travers l'Observatoire de la Grille), combler le fossé entre apprentissage automatique et recherche stochastique, systèmes complexes (évolution artificielle, simulation multi-agents de systèmes sociaux, physique statistique, réseaux neuronaux), optimisation continue (recherche stochastique et algorithmes d'optimisation), aide à la décision optimale sous incertitude (utilisation d'algorithmes de bandits manchots, avec en particulier les résultats spectaculaires du logiciel MoGo sur le jeu de Go, premier à avoir battu des joueurs professionnels), « réservoir computing » (grands réseaux neuronaux récurrents à connectivité aléatoire).

Cette dernière équipe commune au LRI et à Inria, à ce jour encore située bâtiment 490 du campus d'Orsay, rejoindra Digiteo Moulon début 2012.

Les quatre équipes de recherche propres LRI (Université Paris-Sud/CNRS) présentes au PCRI sont :

- **Algo : Algorithmique et complexité**

Cette équipe mène des recherches fondamentales en algorithmique, en combinatoire, en méthodes probabilistes, en algorithmes approchés, en logique et complexité, en vérification de modèles, en théorie des jeux.

- **GraphComb : Théorie des graphes et optimisation combinatoire**

Ces informaticiens développent des recherches en théorie des graphes (cycles dans les graphes orientés, décompositions et placements, aspects algébriques, domination), en optimisation combinatoire déterministe (relaxation semi-définie positive, optimisation robuste, algorithmes d'approximation) et en optimisation combinatoire stochastique (résolution de problèmes stochastiques à deux niveaux, modélisation du risque et prise en compte de contraintes probabilistes). Ils appliquent également une partie de ces recherches à la résolution de problèmes pratiques (réseaux, planification, allocation de ressources).

- **ForTesSE : Test formel et exploration de systèmes**

Cette équipe de recherche travaille dans le domaine des méthodes formelles appliquées à la vérification formelle et au test systématique de logiciels. Elle développe des méthodes et outils pour : le test boîte noire à base de preuve à partir de représentations logiques, le test actif et passif à partir de modèles comportementaux, le test boîte blanche en utilisant des méthodes d'exploration aléatoires. Elle s'appuie sur des activités théoriques sur les formalismes (logique, systèmes de transition, algèbres de processus) et les méthodes (preuve de théorèmes, évaluation symbolique, exploration aléatoire de grands modèles) et elle applique ses résultats aux systèmes communicants (protocoles de communication, services Web), aux systèmes critiques embarqués (transports) et aux infrastructures de sécurité. Parmi les outils développés, on peut citer HOL-TestGen, générateur de cas de test construit au dessus de Isabelle/HOL.

- **Réseaux**

Ces travaux de recherche s'appliquent au domaine des réseaux informatiques, tels que les réseaux sans fil et mobiles, les réseaux ad hoc, les réseaux de capteurs et les réseaux fixes. Les problématiques traitées concernent la qualité de service, la sécurité, le contrôle d'admission, la réservation de ressources, le calcul de bornes déterministes et la gestion de la consommation d'énergie. Pour apporter des solutions à ces problèmes, la méthodologie utilisée consiste à proposer de l'algorithmique qui prend en considération les contraintes d'implémentation dans les réseaux existants ou ceux du futur. L'analyse de performance se fait théoriquement pour déterminer les limites de fonctionnement puis par simulation pour enrichir l'analyse et y inclure la prise en charge d'un environnement dynamique. Cette équipe de recherche propose également plusieurs implémentations de ces solutions sur des systèmes réels. Elle a su concrétiser de longues années d'expertises pour développer une plate-forme de simulation ainsi qu'une implémentation en logiciel libre d'un réseau ad hoc (de la couche physique à l'application).

L'équipe propre Inria est :

- **AVIZ : Analyse visuelle**

L'objectif de cette équipe pluridisciplinaire est d'améliorer les méthodes d'analyse et de visualisation de grandes quantités de données en intégrant profondément le processus d'analyse et celui de visualisation d'information pour permettre de comprendre plus facilement et rapidement ces données. Ses axes de recherche sont : méthodes pour visualiser et naviguer dans des masses de données ; méthodes pour analyser et réduire efficacement des masses de données afin de les rendre visualisables ; méthodes d'évaluation pour mesurer l'efficacité et l'utilisabilité des visualisations, navigations et analyses ; outils logiciels pour réaliser et déployer des systèmes d'analyse visuelle pouvant gérer, chercher, visualiser et analyser des masses de données avec des temps de réponse interactifs. Elle se concentre sur la visualisation de très gros réseaux (de l'ordre de millions de sommets et d'arêtes) et des séries temporelles (plusieurs milliards d'enregistrements capturés en continu et en temps réel). Les domaines d'application incluent l'analyse de grands réseaux sociaux (Wikipedia, les développeurs de logiciels libres), les réseaux biologiques, l'intelligence économique, les bibliothèques numériques et les séries temporelles issues des activités de chercheurs.

Quelques mots d'histoire... et d'actualité

A sa création en janvier 2002, le PCRI devait être une structure de recherche regroupant le LRI, l'Inria Futurs et le LIX (laboratoire d'informatique de l'Ecole Polytechnique).

Le PCRI a préfiguré la création en 2006 du RTRA Digiteo, dans lequel il s'est fondu, et qui rassemble plus de 2000 chercheurs en STIC. L'Université Paris-Sud, le CNRS et Inria sont trois des six membres fondateurs de Digiteo ; et les équipes LRI (Université Paris-Sud/CNRS) et Inria du PCRI sont depuis le début très actives dans cette structure.

DIGITEO

Depuis l'origine, c'est un professeur de l'Université Paris-Sud au sein du LRI qui assure la présidence du comité des programmes de Digiteo, une chaire Digiteo a été attribuée en 2009 à un professeur de l'Université Paris-Sud au sein du LRI, les membres du PCRI ont eu plusieurs projets acceptés dans le cadre du DIM Logiciels et Systèmes Complexes de la région Île-de-France avec un co-financement de Digiteo et plusieurs Opérations de Maturation Technico-Economiques ont vu le jour, l'une venant de déboucher sur la création d'une start-up, MetaScale, sur le HPC, créée par un membre LRI de l'équipe Parall (une première start-up, Green Communications, dans le domaine des réseaux, avait été co-fondée récemment par un membre LRI de l'équipe Réseaux).

System@tic

Les liens des membres du PCRI avec le pôle de compétitivité System@tic Paris-Région sont également étroits. Des membres LRI et Inria du PCRI participent aussi aux groupes programmatiques de l'Alliance des Sciences et Technologies du Numérique, Allistène.

ICT Labs

Le PCRI est également très présent au sein d'ICT Labs, une des trois Communautés de Connaissance et d'Innovation créées fin 2009 par l'Institut Européen d'Innovation et de Technologie, réunissant cinq pays européens. Paris-Sud et Inria font partie des six partenaires français et le nœud français est Paris-Saclay. Un professeur de l'Université Paris-Sud au sein du LRI est responsable au niveau européen d'un des six thèmes d'ICT Labs : Villes numériques du futur. Les équipes Réseaux et BD-IASI/Leo sont fortement impliquées dans les actions recherche et deux projets de master européens coordonnés par des équipes du PCRI sont en cours de montage sur l'Interaction et la Conception Homme-Machine et sur les Systèmes et Services Distribués et l'Informatique en Nuage.

Investissements d'avenir

Les équipes du PCRI se sont mobilisées pour répondre aux appels à projets dans le cadre des investissements d'avenir.

- L'Equipex Digiscope, dont le responsable est un professeur de l'Université Paris-Sud au sein du LRI, a été accepté en janvier 2011 et reçoit une dotation de 6,7M€ pour un budget total de 22M€. Les partenaires sont : la FCS Campus Paris-Saclay, l'Université Paris-Sud, le CNRS, le CEA, Inria, l'Institut Telecom, l'Ecole Centrale, l'Université Versailles-Saint-Quentin, l'ENS Cachan et la Maison de la Simulation. Les équipes In Situ et AVIZ du PCRI sont fortement impliquées dans ce projet. Cet équipement est un réseau de neuf plateformes de visualisation interactive haute performance dont fait partie la plateforme WILD visitée aujourd'hui. D'ici trois ans, ce réseau permettra à des spécialistes de travailler en collaboration et à distance dans quatre grands domaines nécessitant la manipulation de données complexes : la découverte scientifique, la conception de produits nouveaux, l'aide à la décision, notamment en situation de crise, ainsi que la formation et l'enseignement. Digiscope sera utilisé par les partenaires du projet mais également accessible aux utilisateurs extérieurs, notamment industriels. Il a également vocation à s'étendre à d'autres plateformes, en France et à l'étranger.
- Le projet de Labex en STIC DigiWorlds, Mondes numériques : données, programmes et architectures distribuées, coordonné par un professeur de l'Université Paris-Sud au sein du LRI, a été soumis le 12 octobre 2011. Il réunit environ 350 chercheurs de 14 des principales institutions partenaires de Digiteo et toutes les équipes du PCRI y sont représentées. Trois axes de recherche ont été identifiés : SciLex (sûreté du logiciel et sécurité des données), ComEx (systèmes numériques passant à l'échelle, en termes d'architectures et réseaux distribués) et DataSense (gestion de grandes masses de données), qui est coordonné par trois membres du PCRI, responsables des équipes A&O/TAO, BD-IASI/Leo et In Situ. S'y ajoutent des actions de recherche pluridisciplinaires dans le cadre de l'Idex (médecine-santé, biologie systémique, société de l'information, ingénierie). Les enseignants-chercheurs rattachés au LRI ont largement participé à la définition des actions de formation, au niveau Licence, Master et Doctorat.
- Plusieurs équipes du PCRI sont aussi impliquées dans l'IRT SystemX, Ingénierie numérique des systèmes du futur, classé 7^{ème} et recommandé avec examen spécifique. Le projet, avec un budget restreint à 150M€ et une focalisation sur deux programmes (Ingénierie et infrastructure pour le numérique, Systèmes de systèmes) est en cours de finalisation.
- Enfin, plusieurs membres du PCRI se sont impliqués dans le dossier de l'Idex Paris-Saclay soumis le 20 septembre 2011 et présélectionné le 7 novembre 2011, tant dans les actions de



recherche (les STIC sont l'un des 6 champs disciplinaires de l'Idex) que de formation. Cette Idex préfigure la future Université Paris-Saclay pour 2014.

Les équipes du PCRI, tant côté LRI (Université Paris-Sud/CNRS) que côté Inria, jouent donc, par l'excellence de leurs recherches et leur activité et volontarisme au sein des nombreuses structures locales, un rôle moteur dans l'affirmation de la place de l'informatique sur le Campus Paris-Saclay.

PCRI, le 9 novembre 2011
Philippe Dague,
Professeur à l'Université Paris-Sud,
Directeur du LRI