

**Nom de l'école doctorale** : Ecole Doctorale des Sciences Fondamentales de l'Université Blaise Pascal

**Numéro de l'école doctorale** : ED178

**Nom de l'unité de recherche 1** : Laboratoire de Physique Corpusculaire

**Label de l'unité de recherche 1** : UMR6533 (CNRS – Université Blaise Pascal)

**Equipe interne 1** : Plate-Forme de Calculs pour les Sciences du Vivant  
(<http://clrwww.in2p3.fr/PCSV>)

**Localisation** : Aubière - France

**Nom du directeur de thèse** : Vincent Breton

**Nom de l'unité de recherche 1** Laboratoire MSI : Modélisation et Simulation Informatique des systèmes complexes

**Nom de l'établissement** : IFI (Institut de la Francophonie pour l'Informatique)

**Localisation** : Hanoi - Vietnam

**Nom du Co-directeur de thèse** : Nguyen Hong Quang

**Adresse courriel des contacts scientifiques** :

[breton@clermont.in2p3.fr](mailto:breton@clermont.in2p3.fr),

[nguyen.hong.quang@auf.org](mailto:nguyen.hong.quang@auf.org)

**Titre de la thèse** : Etude du déploiement d'un réseau sentinelle pour la surveillance de la grippe aviaire avec la technologie des grilles

### **Description du projet**

Les maladies émergentes constituent une menace croissante pour la santé publique. L'accroissement permanent des échanges internationaux lié à la mondialisation du commerce et au développement du tourisme fait qu'un nouveau virus apparu à n'importe quel point de la planète peut se retrouver en moins de 24 heures à l'aéroport Roissy- Charles de Gaulle. Les maladies émergentes sont caractérisées par le fait qu'elles sont par définition peu connues et qu'il est essentiel de pouvoir accumuler rapidement un ensemble d'informations biologiques, épidémiologiques et géographiques sur les foyers identifiés. Ces informations sont souvent très éparpillées géographiquement mais elles doivent être diffusées le plus rapidement possible au sein de la communauté internationale pour alimenter les travaux des chercheurs et pour permettre aux états de prendre des mesures concertées de prévention et de surveillance.

Depuis 10 ans, la technologie des grilles informatiques a permis le développement de véritables infrastructures distribuées fournissant de vastes ressources de calculs et de stockage aux communautés scientifiques. Depuis 2 ans, elle met aussi à la disposition de ses utilisateurs des outils de gestion de données distribuées permettant la création de véritables fédérations de bases de données d'accès sécurisé dont les informations peuvent être exploitées à des fins d'analyses statistiques. Grâce à ces progrès récents, les grilles constituent aujourd'hui des infrastructures très prometteuses pour la mise en place de systèmes de surveillance à l'échelle internationale.

L'objet de la thèse est de développer et de valider un réseau sentinelle de bases de données sur les maladies émergentes avec comme premier cas d'application la grippe aviaire. A partir des informations locales collectées dans les différents pays, des informations épidémiologiques fournies par les organisations internationales (OMS, FAO) et des informations des bases de données de biologie moléculaire telles que GenBank ou l'Influenza Sequence Data Base à Los Alamos, il s'agit de concevoir une fédération de base de données dont les clients seront les acteurs de la surveillance de cette maladie : médecins épidémiologistes, biologistes, vétérinaires,....

Le choix de la technologie s'appuiera sur l'expérience acquise par l'équipe de recherche PCSV qui depuis sa création en 2001 développe et déploie des applications biomédicales sur des infrastructures de grille. L'équipe coordonne l'initiative WISDOM de recherche *in silico* de nouveaux médicaments contre la malaria et la grippe aviaire. Depuis 2 ans, l'équipe s'investit dans la conception d'une plate-forme de gestion sécurisée de données médicales dans le cadre des projets européens EGEE (FP6) et LifeGrid (PRAI) en s'appuyant sur le catalogue de métadonnées AMGA de l'intergiciel gLite développé dans le cadre d'EGEE.

Sur la base de l'expérience acquise par l'équipe PCSV dans le développement d'applications biomédicales sur les grilles, le doctorant concevra l'architecture et prendra en charge le déploiement d'un démonstrateur de réseau sentinelle pour les maladies émergentes.

La première année de la thèse sera consacrée à une analyse des besoins, à la définition d'un modèle de données et au déploiement d'un premier démonstrateur dans plusieurs pays. Le démonstrateur sera déployé à une échelle internationale grâce à un partenariat établi avec des laboratoires asiatiques dans le cadre de trois Laboratoires Internationaux Associés avec la Chine (LIA FCPPL), la Corée (LIA FKPPL) et le Vietnam (LIA FVPPL). L'analyse des besoins des chercheurs permettra notamment de clarifier les conditions de partage des informations biologiques et médicales. Les données seront collectées auprès des agences internationales, dans les banques de données de biologie moléculaire et auprès des établissements en Asie qui collaborent au projet.

La deuxième année sera consacrée à l'amélioration du démonstrateur et au développement de modèles de surveillance ou d'alerte. L'amélioration portera en particulier sur la mise en place de mécanismes de sécurisation du partage des informations et l'enrichissement progressif du contenu scientifique de la base. A partir des données collectées, des algorithmes de déclenchement d'alerte seront étudiés en collaboration avec des médecins épidémiologistes et développés en partenariat avec des laboratoires d'informatique au Vietnam (Laboratoire MSI - Institut de la Francophonie pour l'Informatique d'Hanoï et Institut des technologies de l'Information de l'Académie des Sciences) où les maladies émergentes constituent une menace permanente sur la santé publique.

Au cours de la troisième année, le doctorant poursuivra le développement des services opérationnels sur la fédération de bases de données qui seront validés avec des acteurs de la surveillance de la grippe aviaire.

Le laboratoire MSI (Modélisation et Simulation Informatique des systèmes complexes) de l'Institut de la Francophonie pour l'Informatique d'Hanoï dispose d'un noeud de la grille ACGRID. Trois autres noeuds de cette même grille (IPH, IOIT et CNRS-Maison de la Science) sont ou seront installés à Hanoï. Ils permettront de simuler les scénarios potentiels afin de valider le démonstrateur. MSI développe actuellement une plateforme de simulation appelée **GAMA (GIS & Agent-based Modelling Architecture)**, une plateforme

générique couplée à un système d'information géographique (SIG) pour la modélisation et la simulation orientée agent. Elle a été utilisée pour l'étude et modélisation de la grippe aviaire au Nord Vietnam. MSI a établi des relations avec le CIRAD et l'Institut de recherche vétérinaire qui sont des partenaires thématiques capables de fournir des données épidémiques ainsi que l'interprétation des résultats observés sur le démonstrateur. Le doctorant travaillera en alternance entre Aubière (PCSV) et Hanoi (MSI-IFI).

**A contacter:**

M. NGUYEN Hong Quang  
email: [nguyen.hong.quang@auf.org](mailto:nguyen.hong.quang@auf.org)