

# Évaluation d'applications distribuées : de l'expérimentation aux méthodes formelles

Martin Quinson: candidature EPI AlGorille

Nancy-Université

29 avril 2010

# Évaluation d'applications distribuées

## Comment étudier une application distribuée

- ▶ Approche théorique : Étude **mathématique** [des algorithmes]
- ▶ Science expérimentale : Étude sur un **instrument scientifique** [des applications]
- ▶ Computational science : **Simulation** [d'un modèle du système] in silicio

## Approche théorique

😊 Meilleure compréhension, théorèmes d'impossibilité ; 😞 Tout NP-dur

## Expérimentations : Applications **réelles** sur un plate-forme **réelle**

😊 Crédibilité ; 😞 Difficile et long. Contrôle expérimental ? Reproductibilité ?

## Émulation : Applications **réelles** sur plate-forme **synthétique** (in vitro)

😊 Meilleur contrôle expérimental ; 😞 Encore plus difficile

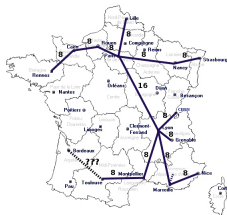
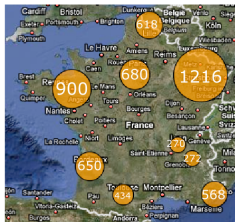
## Simulation : **Prototype** d'applications sur un **modèle** de système

😊 Simple ; 😞 Biais expérimental

# Mes travaux sur l'approche expérimentale

## Projet Grid'5000 : un instrument scientifique pour le HPC

- ▶ Instrument pour la recherche en informatique (*déploiement* d'OS customisés)  
1500 nœuds (3000 cpus, 4000 cores), 9 sites; réseau dédié 10Gb
- ▶ Projet fédérateur de NumB, jalon du plan stratégique INRIA 2008-2012



Experimental conditions injector	Application	Measurement tools
	Programming Environments	
	Application Runtime	
	Grid or P2P Middleware	
	Operating System	
	Networking	

## Mon apport personnel

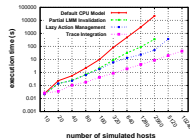
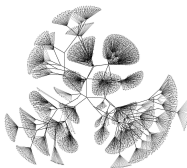
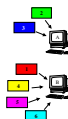
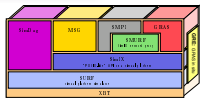
- ▶ En national : **Comité de pilotage**; Animation d'un axe d'**Hémera**; École d'été
- ▶ En local : Projet du **CPER** lorrain; **Animation** scientifique d'utilisation
- ▶ Nancy est un **site moteur** (et attractif)
- ▶ **Collaboration** : Institut des Grilles, **CEA**, Arcelor-Mittal

# Mes travaux en simulation d'applications réparties

SimGrid : simulateur de **prototype** d'applications sur des **modèles** de plates-formes

- ▶ Scalable (temps et mémoire) ; Modulaire (modèles TCP et autres) ; Portable
- ▶ +70 publications ; 100 membres sur simgrid-user@

SIM GRID



Mon apport personnel (en collaboration avec A. Legrand et F. Suter)

- ▶ Principal **architecte logiciel**
- ▶ **Porteur du projet ANR fédérateur** (19 partenaires dans 7 laboratoires)  
Simulation parallèle, Validation applicative MPI, modèles énergie
- ▶ **Collaborations** : CERN, U. Anvers, COST énergie ; CIFRE à venir (AdHoc Intl.)
- ▶ **Dissémination** : SimGrid User Days 2010 ; Tutoriaux CCGrid CeCMG

# Mes travaux en méthodes formelles

## Model-Checking

- ▶ Objectif : Développer des logiciels sûrs
- ▶ Principe : Étant donnée une situation initiale (“3 nœuds interagissent”), tester toutes les exécutions possibles (ordonnancement des messages)
- ☹ Explosion combinatoire

## Mon apport personnel (en collaboration avec S. Merz)

- ▶ Distribué → Model-checking
  - ▶ Ajout de model-checking à SimGrid (thèse 2010-2012)
  - ▶ Prototype efficace utilisant les similarités entre simulation et model-checking
- ▶ Distribué ← Model-checking
  - ▶ Étendre l’outil PlusCal pour exprimer des algos distribués (thèse 2009-2012)
  - ▶ Collaboration avec L. Lamport

# Mon projet de recherche

## De l'expérimentation aux méthodes formelles

- ▶ **Réflexion méthodologique** sur les problèmes spécifiques de chaque approche  
Contrôle, reproductibilité, biais expérimental, applicabilité ...
- ▶ **Réalisation d'outils pratiques** : Grid'5000, SimGrid, PlusCal
- ▶ **Approche transversale** : à problèmes communs, solutions communes

## Réflexions méthodologiques transversales Grid'5000/SimGrid

- ▶ **Open Science** : cahiers de laboratoire et workflow d'expériences  
Le graal de la reproductibilité ; la norme en physique, chimie, biologie, médecine, ...
- ▶ **Plans d'expériences fractionnaires**  
Rationalisation des tests, réduction du nombre de tests ; classique en biologie

## Performance-checking

- ▶ De l'étude formelle de la **correction** à celle des **performances**  
sans doute plus intéressant pour le High Performance Computing
- ▶ Gros défi : le temps est absent ou discrétisé en model-checking

# Conclusion : mes apports potentiels

## Au centre INRIA Nancy Grand Est

- ▶ Pérénisation de l'axe **méthodologie expérimentale** de l'EPI AlGorille
- ▶ Renforcement du **pont thématique** entre les EPI AlGorille et Veridis
- ▶ **Rayonnement** accru au travers des projets Grid'5000 et SimGrid

## À l'INRIA

- ▶ Méthodologie d'expérimentation d'applications distribuées : **EPI à venir ?**
- ▶ Expertise et animation autour des **grilles expérimentales**
- ▶ Expérience en responsabilité et **coordination de contrats de recherche**
- ▶ **Nombreuses collaborations** académiques et industrielles

## Impact à ce jour

- ▶ 1 chapitre de livre, 4 journaux, 14 conférences, 3 communications courtes
- ▶ 5 logiciels (250 000 lignes), dont 3 largement utilisés (centaines d'utilisateurs)
- ▶ Porteur de 8 projets dont axe CPER, ANR ARPEGE (1,5M€ depuis 2005)
- ▶ Membre du comité de pilotage Grid'5000, chargé de mission CRI NGE

# Agenda

- Curriculum Vitæ
- Mes travaux de recherche : Évaluation d'applications distribuées
  - Approche expérimentale
  - Approche par simulation
  - Approche formelle
- Mon projet : de l'expérimentation aux méthodes formelles
- Conclusion