Statistical Model Checking in a nutshell

Benoît Delahaye

Université de Nantes

2014-02-06

Why do we need probabilities ?

Principalement pour faire de l'abstraction

- Phénomènes complexes
  - Trop complexes à modéliser
  - Dont on ne connaît pas le fonctionnement précis
- Environnement inconnu
- ...

Mais aussi : Algorithmes probabilistes...

Probabilistic Modeling...

Beaucoup de formalismes

- Automates
- Logiques
- Algèbres de processus
- Systèmes hybrides
- ...

... is more complex than expected

- Non-déterminisme : Quel sens lui donner ?
- Concurrence
- Temps
- Ressources / autre notions quantitatives

What to do with probabilistic models/systems ?

- Simulation
- Test
- Vérification

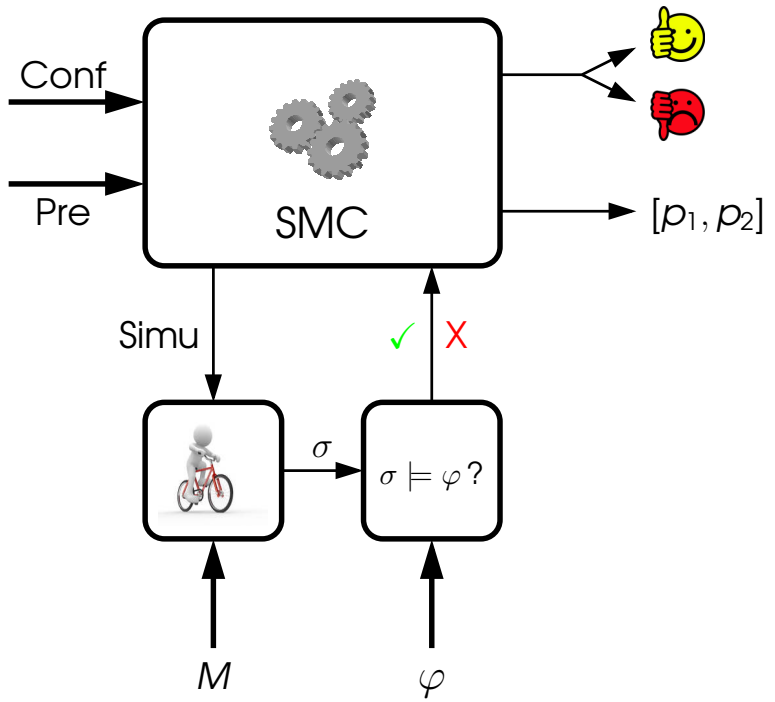
Outline

Outline

Statistical Model Checking

$$\mathbb{P}(M \models \varphi) \geq \alpha\,? \quad \mid \quad \mathbb{P}(M \models \varphi) = ?$$

- Approximation statistique par simulations
- Algorithmes statistiques
  - **Précision garantie**
  - **Probabilité d'erreur garantie**
- Méthodes quantitatives
- Efficacité
- **Estimation de l'indécidable**



Outline

$\mathbb{P}$-BLTL

On ne peut pas vérifier n'importe quelles propriétés avec SMC.

$\Rightarrow$ Propriétés linéaires bornées

Bounded LTL

- **T**, **F**, p, $\neg p$, for all $p \in AP$;
- $\varphi_1 \vee \varphi_2$, $\varphi_1 \wedge \varphi_2$, where φ_1 and φ_2 are BLTL formulas ;
- $\bigcirc \varphi_1$, $\varphi_1 \mathcal{U}^t \varphi_2$, where φ_1 and φ_2 are BLTL formulas, and t is a positive integer.

Probabilistic Satisfaction

Satisfaction of a BLTL formula

Probability for a Markov Chain M to satisfy a BLTL formula φ :

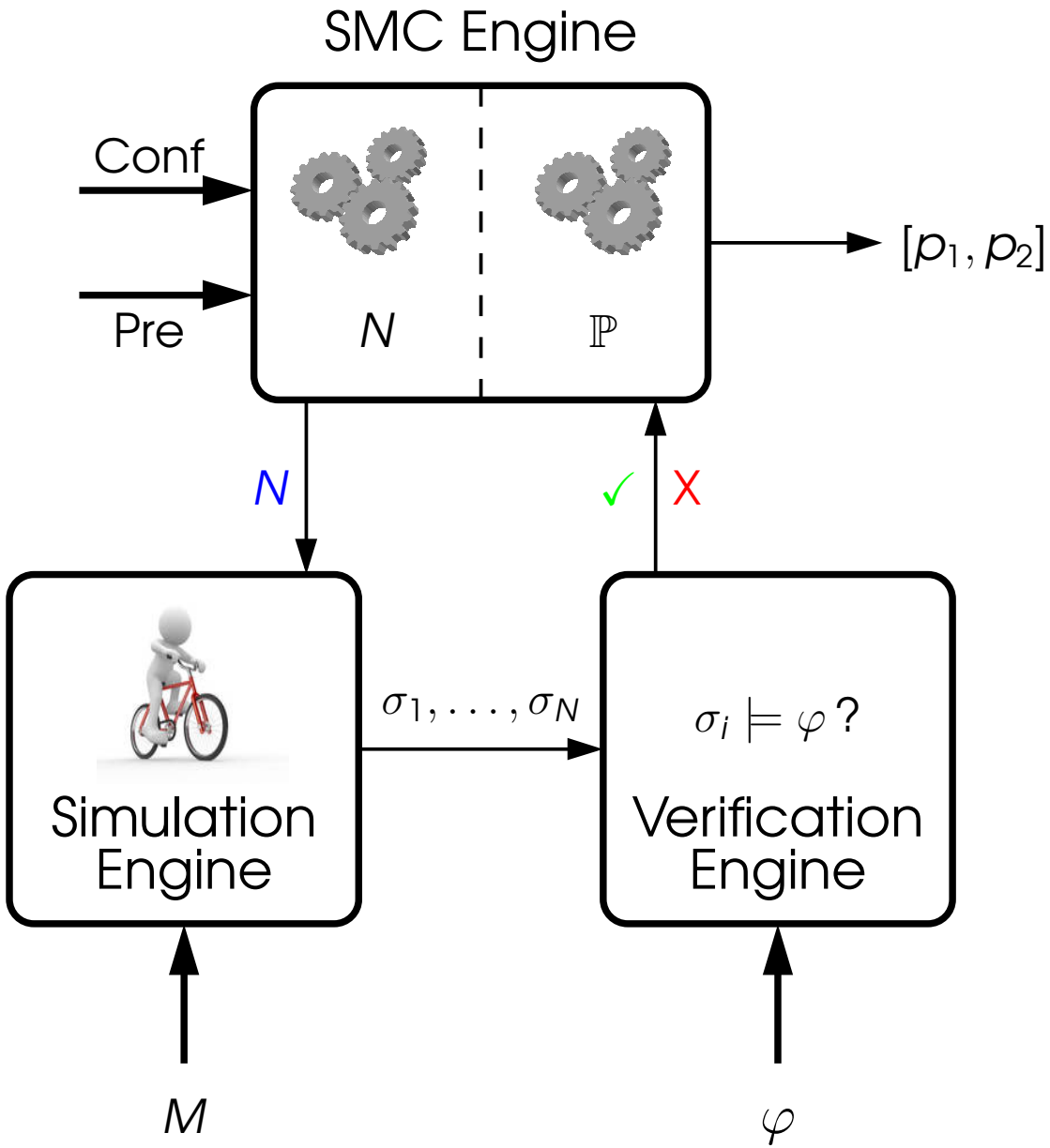
$$\mu\{\pi \mid \pi \models \varphi\}$$

where

- π are executions of M
- μ is the underlying probability measure

Outline

Computing $\mathbb{P}(M \models \varphi)$



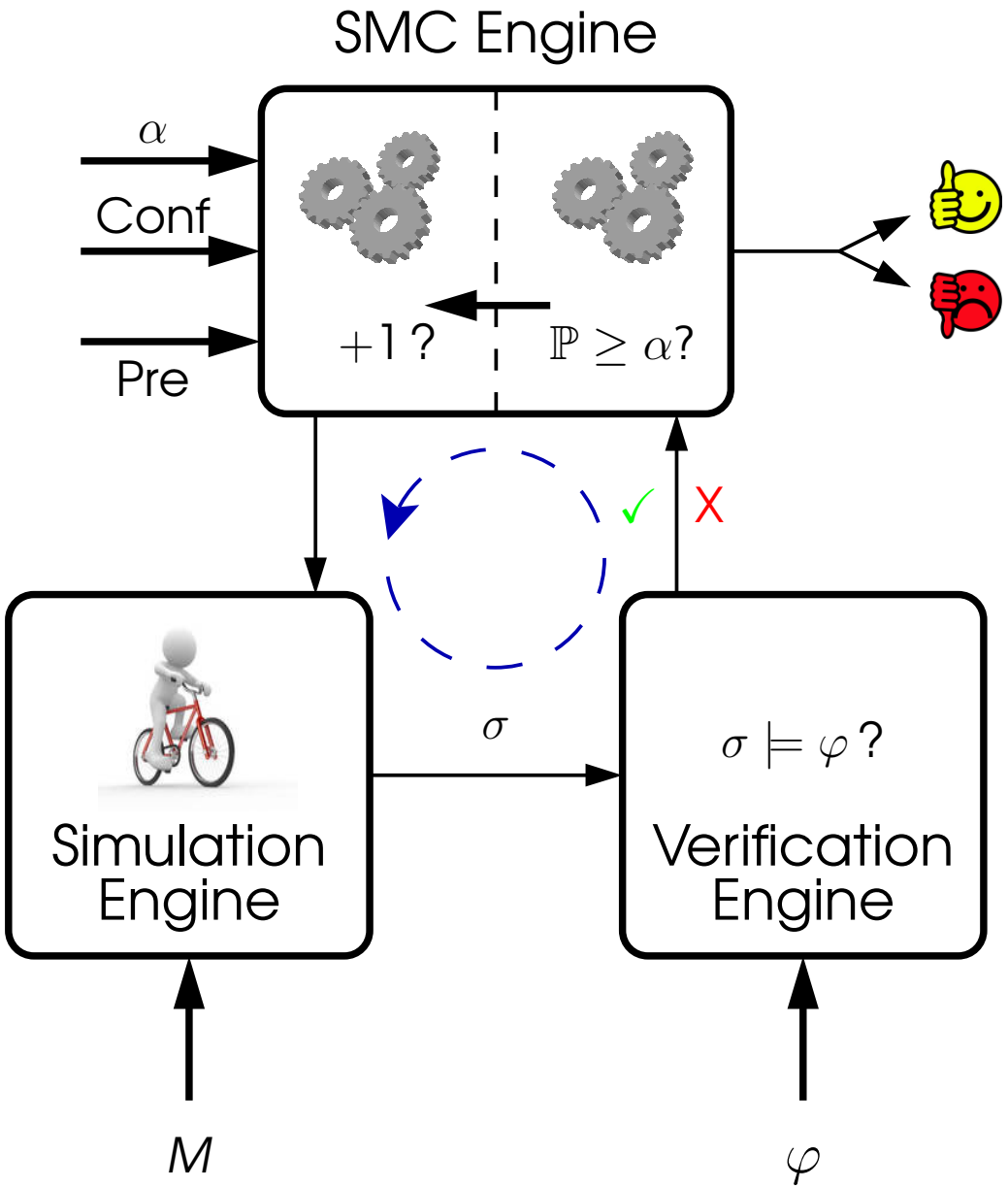
Advantages and Drawbacks

+ Très simple à mettre en oeuvre
+ Pas de boucles, peu de liens entre le calcul de N et le résultat
+ Parallelisation évidente

– En fonction de `pre` et `conf`, N peut être très grand
– M doit être pûrement probabiliste

Outline

$\mathbb{P}(M \models \varphi) \geq \alpha\,?$



Advantages and Drawbacks

+ Très efficace (si bien règlé)

– Nécéssite une connaissance de α

– Plus complexe à mettre en oeuvre

– Plus difficile à paralleliser

– M doit être pûrement probabiliste

Outline

Case Study : Accuracy of clock Synchronization

Chalenges :

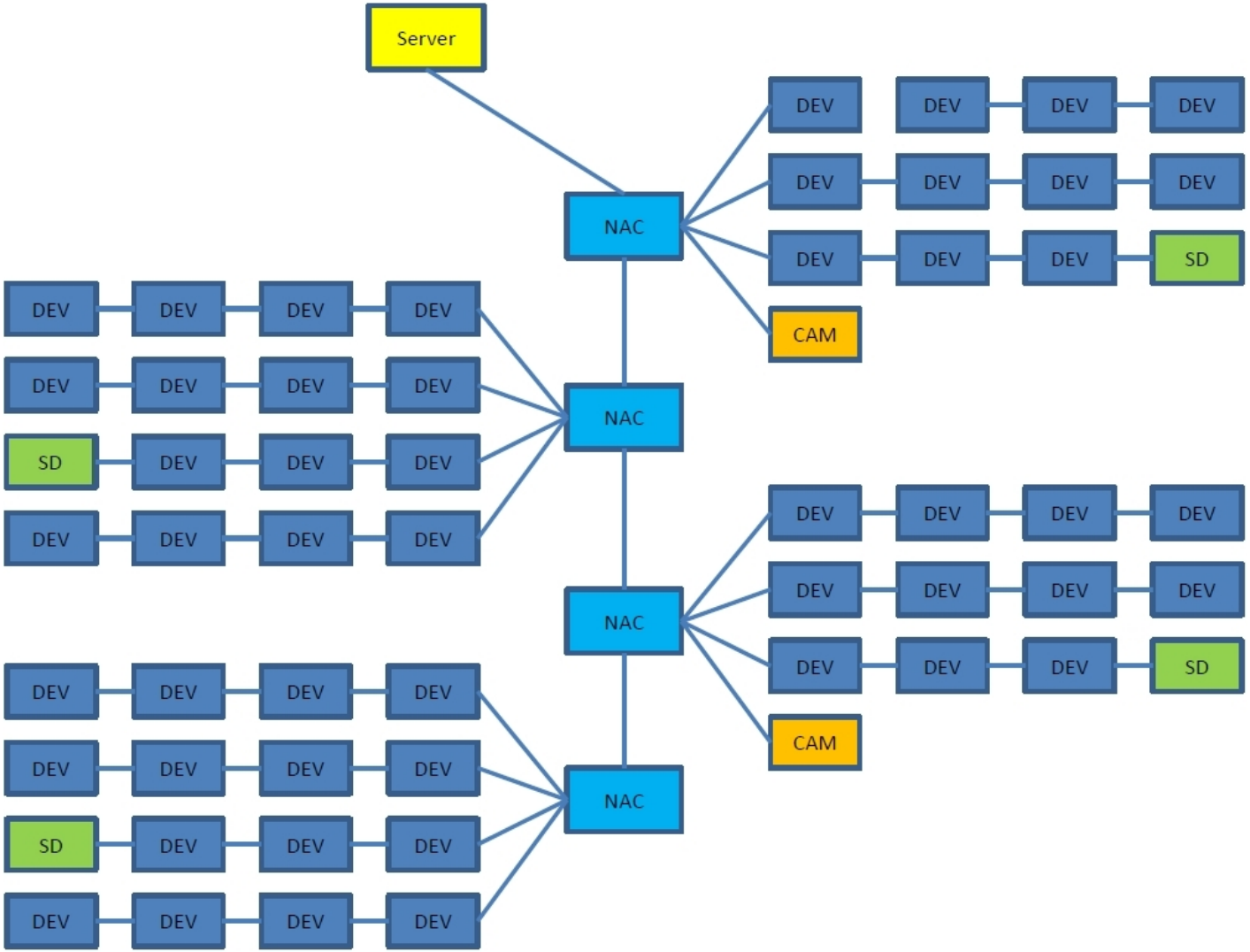
- Heterogeneous System over an Ethernet backbone
  - Distributed application
  - 280 communicating components
- Local clocks synchronized using the Precision Time Protocol
- Requirement : Verify that the difference between any 2 clocks is lower than a given bound

Case Study : Accuracy of clock Synchronization

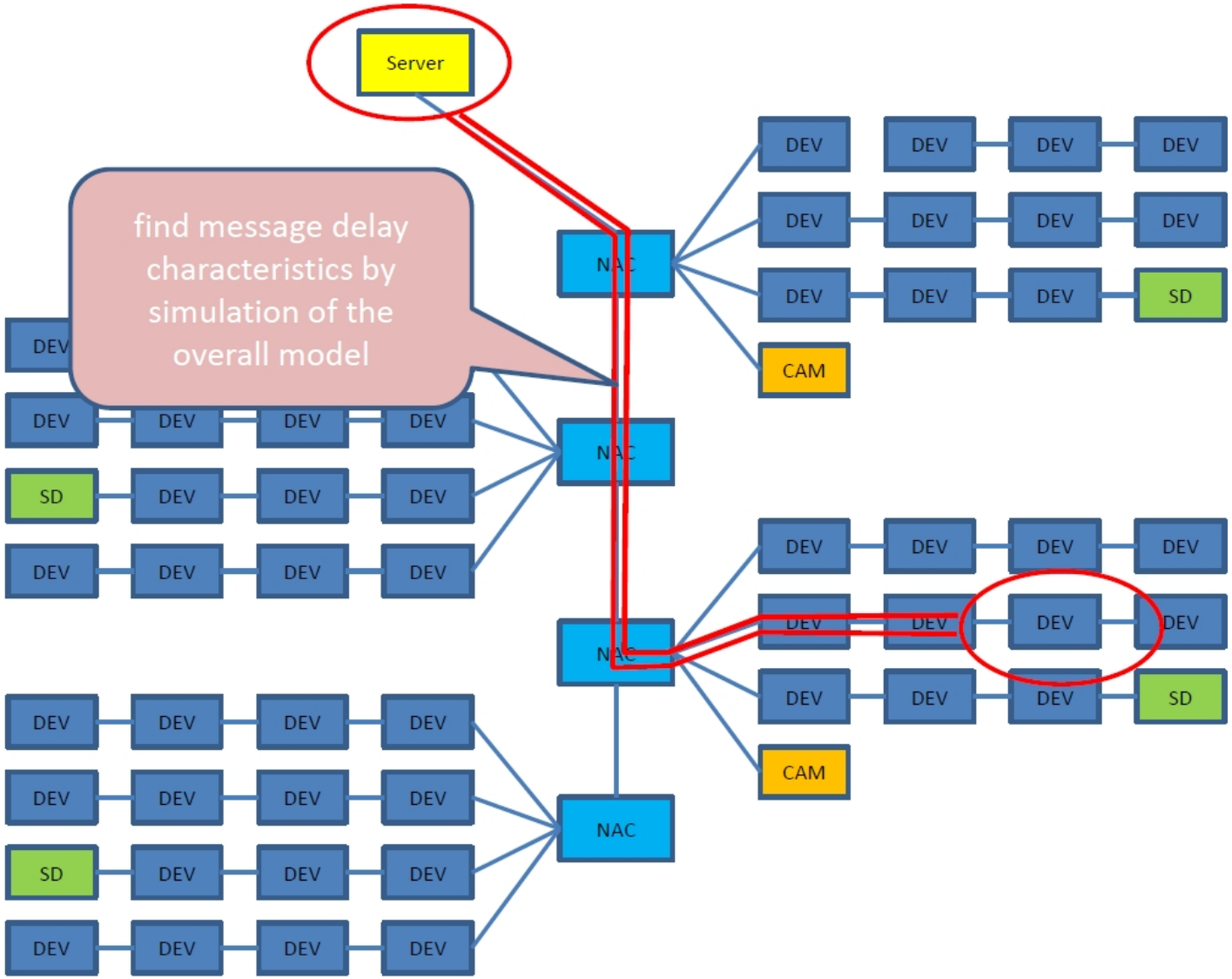
Chalenges :

- Heterogeneous System over an Ethernet backbone
  - Distributed application
  - 280 communicating components
- Local clocks synchronized using the Precision Time Protocol
- Requirement : Verify that the difference between any 2 clocks is lower than a given bound
- Our goal : Compute the best bound to satisfy this requirement without analyzing the whole architecture

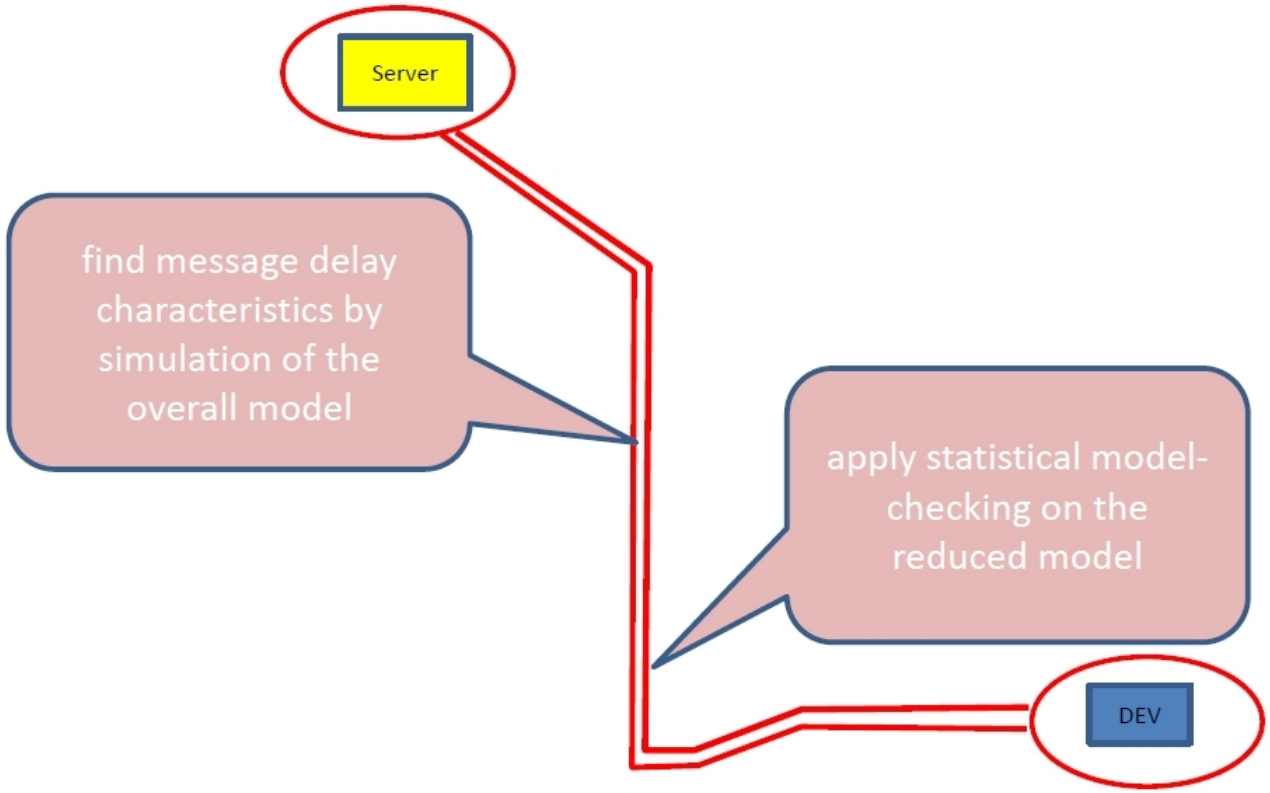
Case Study : Methodology (2)



Case Study : Methodology (2)



Case Study : Methodology (2)

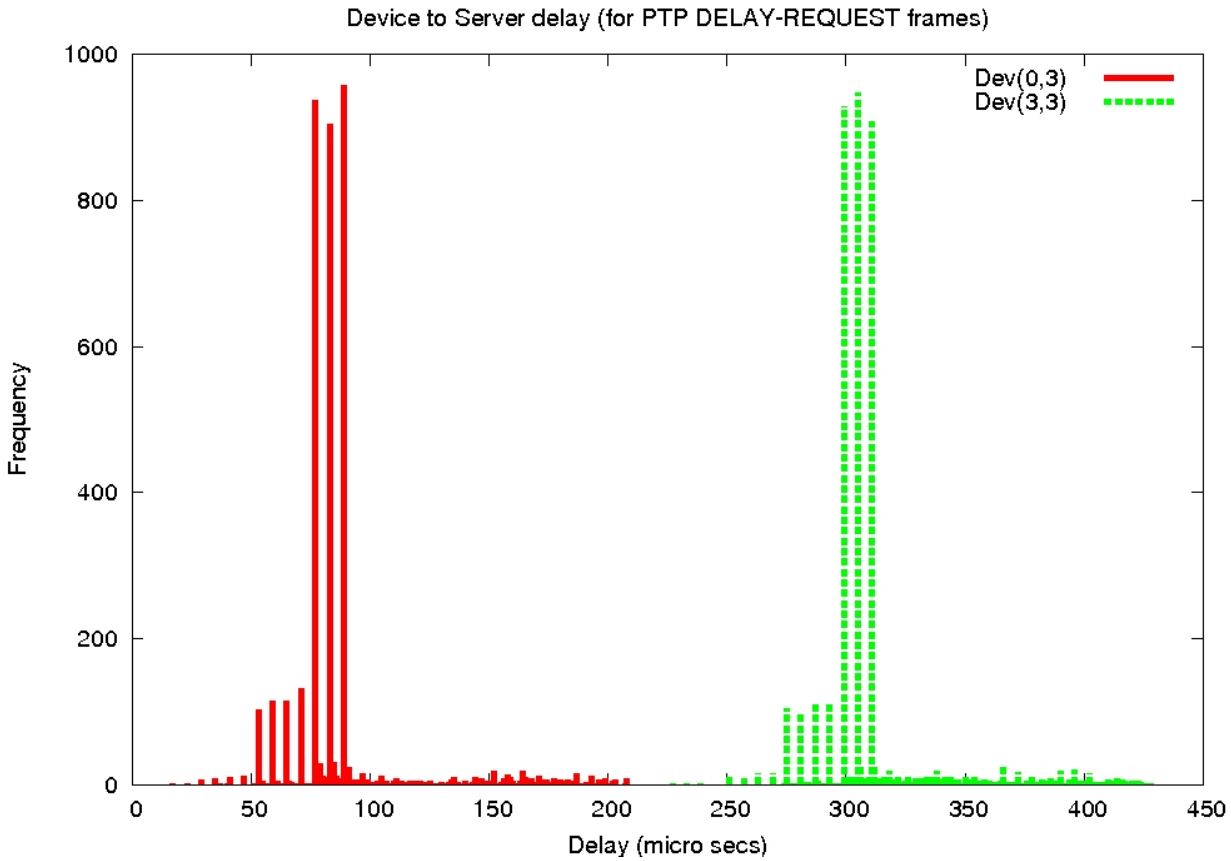


First step

1. Learn the Probability distributions

First step

1. Learn the Probability distributions

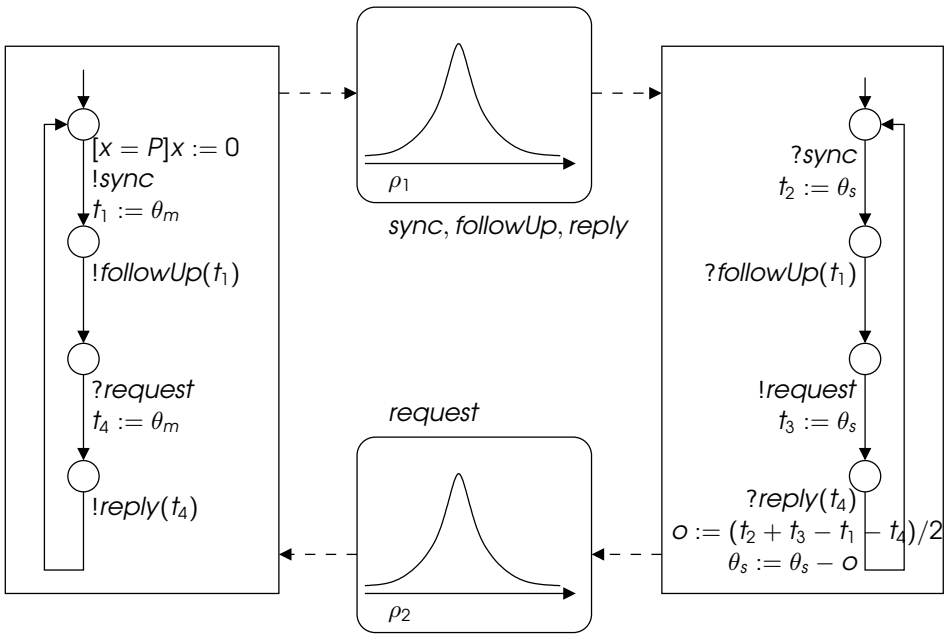


Second Step

2. Use the distributions to study PTP

Second Step

2. Use the distributions to study PTP



In Practice

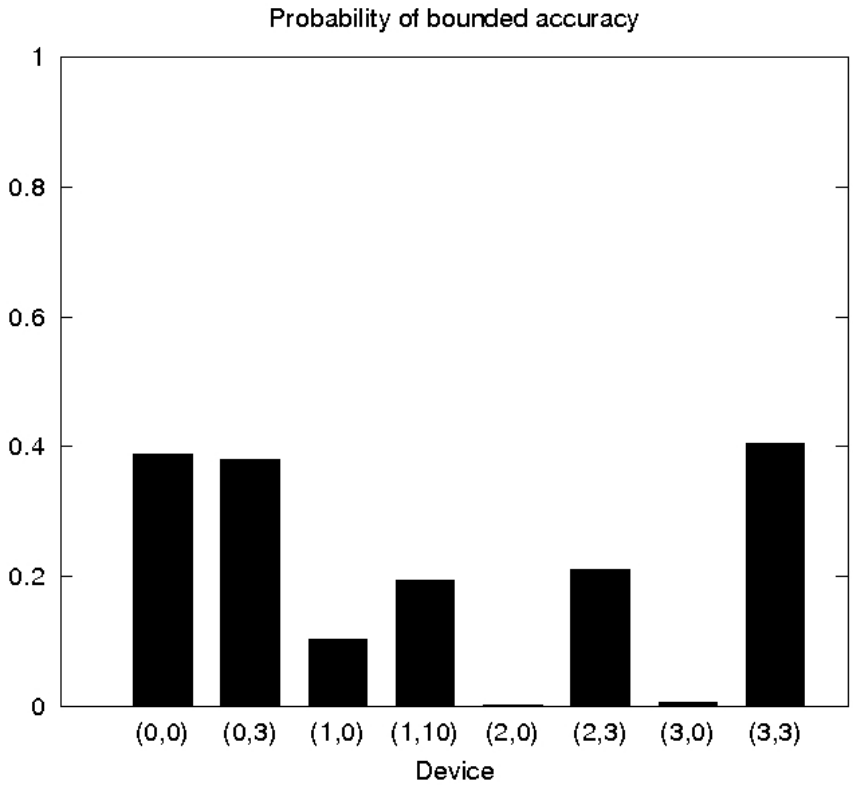
Model/Abstraction :

- PTP and HCS modeled using BIP
- Distributions of delays : 2000 measures

Statistical Model Checking :

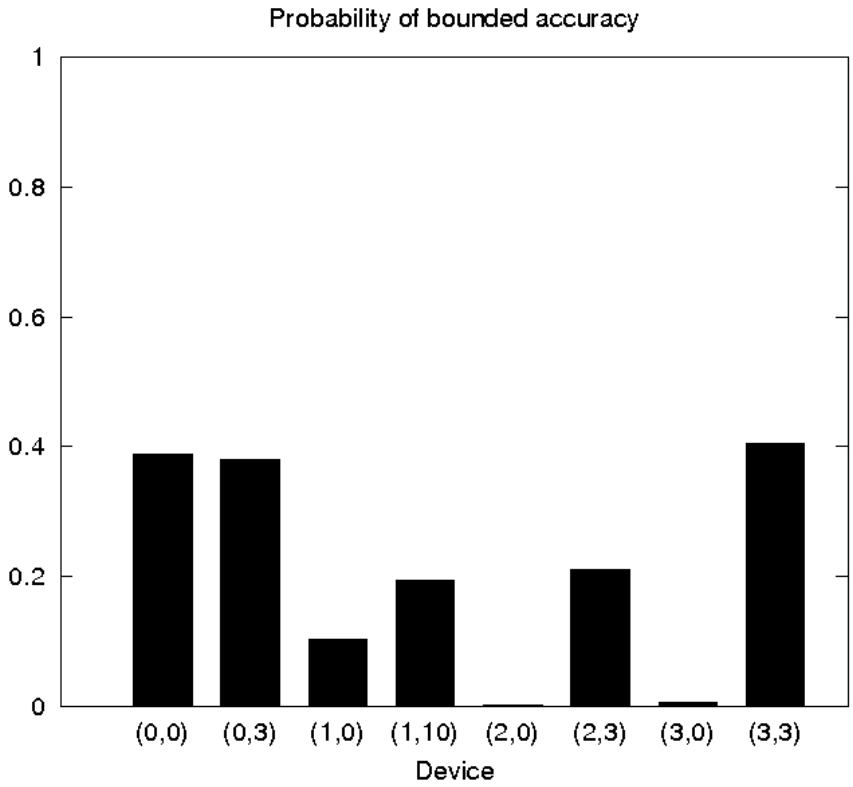
- Quantitative question : precision 10^{-2}, confidence 10^{-2} : 100000 simulations
- Qualitative question : precision 10^{-3}, confidence 10^{-10} : 300000 simulations

Some Results 1/2



Probability of satisfying Bounded Accuracy for a bound of $50\mu s$

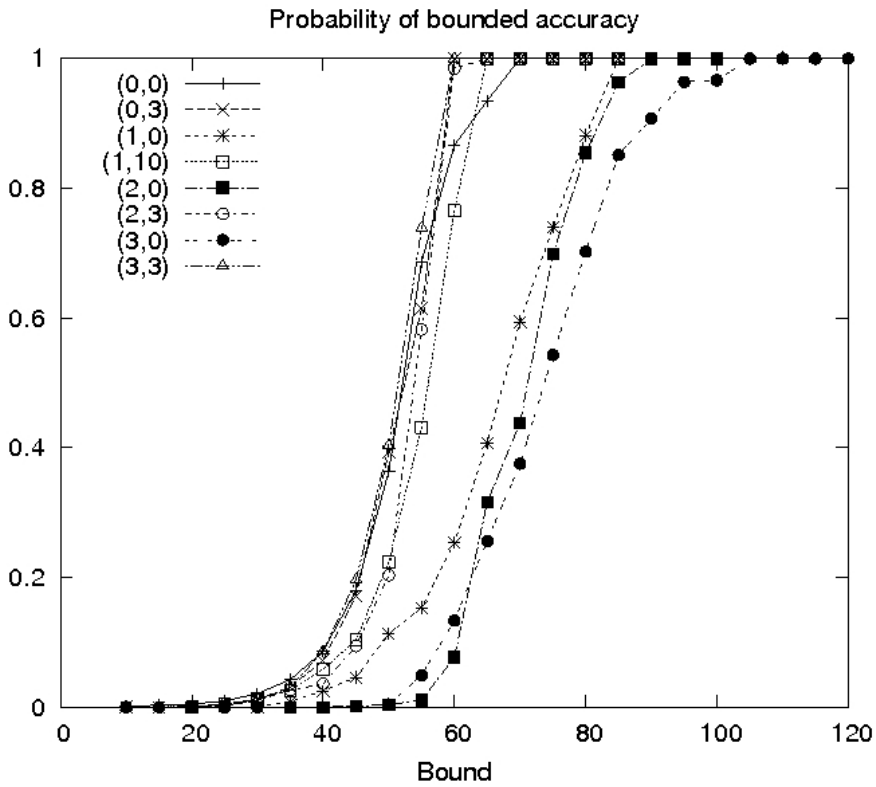
Some Results 1/2



Probability of satisfying Bounded Accuracy for a bound of $50\mu s$

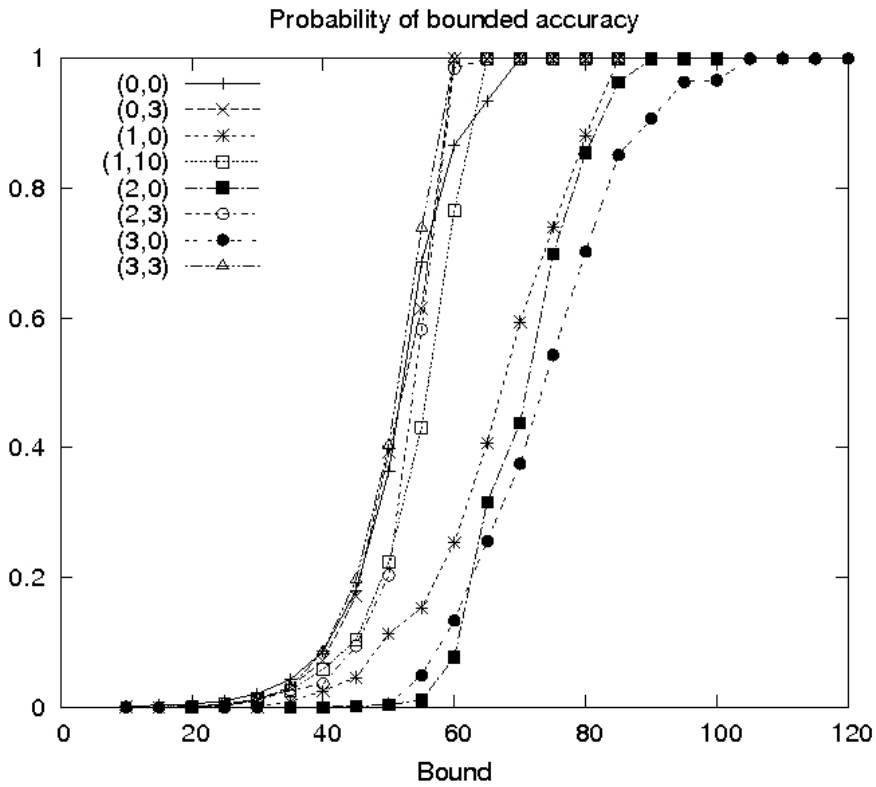
- The property is not satisfied for the given bound !

Some Results 2/2



Probability of satisfying Bounded Accuracy as a function of the bound

Some Results 2/2



Probability of satisfying Bounded Accuracy as a function of the bound

- The best bound for which B.A. is satisfied with probability 1 is $105\mu s$

Conclusion

- Abstraction and verification method
- Applied to 2 case studies :
  - HCS case study [FORTE'10]
  - AFDX network [RV'10]

Outline

Limitations of SMC

Limitation principale : le non-déterminisme

- Systèmes concurrents
- Environnements ouverts

Learning the best scheduler (Henriques'12)

Idée

Utiliser les techniques de learning pour apprendre le scheduler optimal w.r.t φ

Algo :

- début : scheduler uniforme
- apprentissage du scheduler par simulations : N itérations
  - calcul d'une matrice de renforcement des transitions entrainant satisfaction de φ
  - mise à jour du scheduler
- SMC avec la version déterministe du scheduler obtenu

Advantages and drawbacks

+ Algorithme convergent
  vers un optimum local ?
+ Excellents résultats sur certains modèles *organisés*
– Nécéssite de garder en mémoire les matrices de renforcement
– Hypothèse : schedulers sans mémoire
  Insuffisant en pratique

SMC^2

Idée

Utiliser SMC pour trouver le meilleur scheduler.

Algo :

- Générer un scheduler σ aléatoirement
- Calculer la proba de satisfation α de φ par SMC sur le modèle obtenu
- Enregister la paire (σ, φ)
- Recommencer jusqu'à trouver le meilleur σ (ou autre algo)

Advantages and drawbacks

+ Couvre tous les schedulers (en théorie)
+ Permet de calculer d'autres choses que min/max
- Nécéssite une représentation *efficace* des schedulers
- Ne fonctionne pas...

SMC on trees

Idée

Vérifier *tous* les schedulers à la fois.

- Production d'arbres *complets* d'exécutions
- 2 notions de satisfaction :
  1. Arbre OK si *toutes les branches* satisfont φ
  2. Arbre OK si *une branche* satisfait φ
- Résultats : on obtient la une borne inférieure / supérieure sur la probabilité de satisfaire φ

SMC4PN

La concurrence est une forme particulière de non-déterminisme...

Idée

Générer des processus de branchements au lieu des traces.

Difficultés :

- Toutes les traces issues d'un même processus doivent satisfaire les mêmes propriétés
  - Utilisation d'une logique particulière garantissant cette propriété ?
  - Sous-ensemble de RdP ?
- Vérifier les propriétés
  - Directement sur les processus ?
  - Sur une trace de chaque processus ?

Potential applications

- Bio-info
  - Non-déterminisme
  - Formulation des propriétés à vérifier
  - Modèles
  - Temps, Hybride...
- Electronique
- ...

The End

Thank you for your attention