FORMULAIRE de PROPOSITION : Projet École

Date de la proposition :	13 mai 2015

1 – Titre
Modélisation et découverte de systèmes embarqués pour la cybersécurité. Acronyme : MODSOC-Cyber.

2 – Objet
L'étude aborde l'interface entre matériel et logiciel dans un cadre de cyberdéfense, en se focalisant sur la capitalisation dans des modèles les informations découvertes sur les systèmes embarqués. L'objectif général du projet est de constituer une carte d'identité d'un système numérique afin de l'évaluer et ou de le découvrir. Pour se faire, nous nous proposons de découvrir une architecture numérique, à partir de documentation et d'expérimentations, et de constituer et fédérer plusieurs modèles de références de cette architecture.

3 – Titulaire	
Laboratoire - Ecole : Lab-STICC, ENSTA Bretagne	
Adresse : 2, rue F. Verny 29806 BREST Cedex 9	
Coordonnées du responsable du projet	
Nom: Champeau	Qualité : Enseignant chercheur
Prénom : Joël	e-mail : joel.champeau@ensta-bretagne.fr
Tél : 02 98 34 88 42	fax : 02 98 34 89 35

4 – Contact DGA
NOM - Prénom : Frédéric Valette

5 – Vision pluriannuelle : durée & éléments financiers globaux			
Durée (en mois) : 36 (3 ans)			
	Année 2015-2016	Année 2016-2017	Année 2017-2018
Coût marginal	105,8 KE euros HT	105,8 KE euros HT	105,8 KE euros HT

6 – Liens avec d'autres études
Cette étude est en continuité avec les études et activités de l'équipe du laboratoire Lab-STICC de l'ENSTA Bretagne/ Pôle STIC soutenues par la DGA. L'étude MRIS démarrée durant la période 2014/2015, adresse la problématique de gestion de l'obsolescence logicielle/matérielle par virtualisation: il s'agit d'insérer une couche d'abstraction

entre le système réel et la cible vue par l'utilisateur. Ce mécanisme rejoint partiellement la préoccupation de la présente étude ; en revanche, il s'agit ici de produire des modèles du système numérique afin de capitaliser la connaissance sur l'interface entre le matériel et logiciel.

De précédentes études ont également porté sur la modélisation et la validation de modèles systèmes en lien avec le pole SDS, les aspects de modélisation support de ce projet repose également sur ces études puisque nous avons la volonté de modéliser le système à différents niveaux d'abstraction pour augmenter la capitalisation de la connaissance du système.

De plus l'approche de modélisation issue de la thèse (DGA-Région Bretagne) de Jean-Philippe Schneider, axée sur l'interopérabilité de modèles dans un cadre de systèmes de systèmes est reliée avec cette proposition puisque nous devrons faire coopérer différents modèles de manière dynamique en fonction des analyses effectuées sur le système.

Enfin, ces travaux s'articuleront naturellement avec les activités menées dans le cadre du CPER Cyber-SSI, et plus particulièrement du sous projet 4 (Cyber ICS), dans lequel des dispositifs des systèmes SCADA sont émulés sur FPGAs.

7 – Synthèse de la proposition

Le projet a pour objectif d'adresser **les aspects de modélisation de l'interface entre matériel et logiciel dans un contexte de cyberdéfense**.

Dans ce cadre général, le projet a pour but de **constituer une méthodologie outillée de découverte d'information et de capitalisation sur un système embarqué susceptible d'être attaqué**. Les **systèmes embarqués de type SCADA** représentent une bonne illustration pour ce type de préoccupations. En ce sens, la méthodologie peut se présenter comme une activité de reverse-engineering. Dans l'attaque d'un tel système embarqué, l'articulation logicielle-matérielle est cruciale : quels processeurs sont susceptibles d'intervenir dans la conception ? Quels bus ? Quelle est la structure de la memory map ? Quels périphériques? De quelles marques ?
Ces informations sont généralement disparates et incomplètes, mais elles existent. **Notre projet se propose de capitaliser dans des modèles ces informations.**

La capitalisation de ce savoir se fait donc à travers des modèles hétérogènes et initialement incomplets, que l'on raffine au fur et à mesure des découvertes sur ce système. Pour ces modèles constitués par incréments, nous essayerons le plus possible de les conserver exécutables, malgré l'imprécision dans la connaissance du système réel. Par exécutables, nous entendons simulés ou synthétisés sur matériel de type FPGA pour conserver les propriétés du système embarqué, notamment pour les systèmes SCADA qui réclament une exécution temps réel. **L'objectif de ce prototypage FPGA est de confronter au plus tôt les hypothèses prises sur le système.**

La modélisation outillée que nous nous proposons de mettre en œuvre repose **sur une méthodologie agile basée sur des modèles hétérogènes** (différents langages et différents niveaux d'abstraction) allant du niveau système jusqu'à **la modélisation de la plateforme matérielle**. Dès lors, il s'agit de trouver les moyens technologiques de reconstruire au plus tôt un modèle de système, même approximatif. Dès lors, nous pouvons activer une pile logicielle avancée : génération de code, synthèse comportementale, estimation de paramètres, etc… Ainsi nous pourrons simuler, estimer, prototyper une architecture de système embarqué qui nous permettra d'infirmer ou de confirmer les hypothèses capitalisées dans les modèles. **Cette partie du projet est fortement basée sur des compétences d'architectures matérielles et des flots de conception associés.**

La phase de découverte d'information et de capitalisation s'appuie sur une modélisation du système. **Le système embarqué étant par nature multi-points de vue et intégrant différents niveaux d'abstraction, la modélisation du système repose sur une approche de modélisation hétérogène.** Cette modélisation doit donc intégrer différents langages définissant le système avec différents degrés de précision et aussi être capable **d'assurer une fédération entre ces modèles**. Cette fédération de modèles évoluant au cours de la découverte du système, elle doit reposer conceptuellement et technologiquement sur **une approche permettant une évolution dynamique des modèles et de leur fédération au cours du temps.** Cette fédération reposera sur **une modélisation par rôle** qui offre cette capacité de modélisation dynamique de la fédération au cours du processus de modélisation.

La proposition d'étude s'inscrit par conséquent dans une démarche duale, de génie logiciel basé sur les modèles d'une part, et de maitrise technologique du matériel d'autre part. Ainsi, elle mobilise **des compétences issues de plusieurs champs disciplinaires : ingénierie matérielle et ingénierie logicielle**.

8 – Partie scientifique et technique

8.1 Objectifs et enjeux - Problématique

La cyber sécurité est aujourd'hui identifiée comme un enjeu critique. De nombreux travaux ont été initiés sur le volet matériel (cryptographie, résilience, etc.). De nombreuses initiatives s'intéressent au volet logiciel (bonnes pratiques de programmation, etc.). En revanche, très peu de travaux ciblent l'interfaçage entre matériel et logiciel, pourtant présent dans tous les systèmes et notamment les systèmes SCADA. Nous pensons que cette articulation, n'est pas seulement un risque, mais également une opportunité. Renforcer la méthodologie de découverte des connaissances de cette interface matérielle logicielle permet de mieux maitriser les systèmes à sécurisés mais aussi de proposer une méthodologie de découverte de systèmes non connus de manière exhaustive, par exemple en vue d'une attaque.
Manipuler plusieurs niveaux d'abstraction et multi-points de vue dans des modèles permet en effet de renforcer la capitalisation des connaissances de l'interface matériel - logiciel.

La méthodologie est selon nous en parfaite adéquation avec des approches de conception descendante (top-down) classiques, où les spécifications subissent des changements incessants et où les raffinements (introduction de détails architecturaux, etc) correspondent à autant d'explorations architecturales. Ces approches incrémentales sont particulièrement employées dans la conception de SoC (system-on-chip), où 2 à trois niveaux d'abstraction différents ont cours : comportemental, architectural et détaillé [1,2,9]. **Ce projet nous permet donc d'envisager une stratégie de capitalisation des connaissances d'un système, par analogie et transposition, selon une ingénierie (conceptuellement) traditionnelle** : l'ensemble du cycle en V est alors applicable. La phase de remontée de ce cycle permet par exemple de confronter les hypothèses retenues lors de la constitution du modèle synthétique avec le système réel ciblé par une attaque. Cette phase peut par ailleurs constituer une étape d'incrément dans la connaissance du système : ce qu'il est et peut être surtout ce qu'il n'est pas.
Dans ce cadre la synthèse de haut niveau permet la production de circuits matériels à partir de spécification modélisées. La maitrise de cette synthèse de haut niveau dans l'équipe de l'ENSTA Bretagne permet de se reposer sur une expertise éprouvée pour **améliorer la modélisation des systèmes sur différents niveaux d'abstraction**. En effet, l'équipe de l'ENSTA Bretagne a développé en propre et en collaboration avec ses partenaires du Lab-STICC des outils logiciels permettant la synthèse de haut niveau [6,7,8]. La maitrise technologique de la phase de synthèse de haut-niveau est critique. Elle permet d'orienter la synthèse à façon, pour répondre à des problématiques ciblées.

En appui de cette maitrise de l'information de synthèse matérielle, **la phase de découverte de capitalisation de l'information s'appuie sur une modélisation hétérogène du système**. En effet, il existe différents langages de description de matérielle sur différents niveaux d'abstraction mais aussi de modélisation au niveau système. **La fédération de ces différents modèles** reste à ce jour une problématique ouverte sans approche largement adoptée même si différentes initiatives cherchent à combler ce manque [3].
La modélisation par rôle qui repose sur de nombreux travaux tant théoriques que technologiques [4,5] se propose de fournir une alternative **pour la définition d'interfaces adaptables en allouant statiquement et dynamiquement des objets ou éléments de modèles par des rôles**.
Dans le cadre cette étude nous nous proposons donc d'utiliser **ce concept de rôle pour fédérer les différents modèles mis en jeu pour la capitalisation des connaissances des systèmes** que nous cherchons à analyser dans un but de découverte de son architecture. Pour cela, nous nous baserons sur les travaux issus d'une thèse (DGA – Région Bretagne) menée au sein de l'équipe qui a défini et implanter un langage spécialisé pour la définition de rôle dans le but de fédérer différents

modèles systèmes [10,11,12].

L'enjeu de l'étude se positionne au niveau de la gestion de la connaissance de l'interface matériel – logiciel et elle repose sur une **articulation entre les métiers de l'ingénierie matérielle et de l'ingénierie du logiciel.** Les différents modèles hétérogènes fédérés servant de capitalisation de cette connaissance lors de la découverte de systèmes en vue d'une sécurisation ou d'une attaque.

1. Z. J. Jia, A. Núñez, T. Bautista, and A. D. Pimentel. 2014. A two-phase design space exploration strategy for system-level real-time application mapping onto MPSoC. *Microprocess. Microsyst.* 38, 1 (February 2014).
2. Mark Thompson and Andy D. Pimentel. 2013. Exploiting domain knowledge in system-level MPSoC design space exploration. *J. Syst. Archit.* 59, 7 (August 2013), 351-360.
3. M. Seifert, C. Wende, and U. Aßmann, "Anticipating unanticipated tool interoperability using role models," in Proceedings of the First International Workshop on Model-Driven Interoperability. ACM, 2010.
4. F. Steimann, "On the representation of roles in object-oriented and conceptual modelling," Data & Knowledge Engineering, vol. 35, no. 1, pp. 83–106, 2000.
5. T. Kühn, M. Leutäuser, S. Götz, C. Seidl, and U. Aßmann, "A metamodel family for role-based modeling and programming languages," in Software Language Engineering, ser. Lecture Notes in Computer Science. Springer International Publishing, 2014,

Les références 6 à 13 sont la partie relative aux références de l'équipe liées au projet.

8.2 –Intérêt Défense

A travers le livre blanc de la défense et de la sécurité nationale, le ministère de la défense a élevé la cybersécurité au rang de priorité nationale. En effet, les moyens de la défense nationale reposent majoritairement sur des systèmes d'information et des systèmes embarqués interagissant entre eux. Le pole d'excellence en cybersécurité situé en Bretagne s'inscrit dans la démarche de renforcement des compétences du domaine au profit du ministère de la défense. L'ENSTA Bretagne est partie prenante dans ce pole de cybersécurité.

8.3 –Compétences et expériences du laboratoire

Le pôle STIC de l'ENSTA Bretagne travaille depuis plusieurs années dans le domaine de la modélisation et validation de systèmes embarqués. L'équipe IDM du pôle a développé des travaux en partenariats avec de nombreux académiques (INRIA, TELECOM BRETAGNE, ONERA, IRIT) et industriels (THALES, AIRBUS, NEXTER, CS-SI) dans le cadre de projets ANR (DOMINO, MOPCOM-Soc), DGE (TOPCASED, MOPCOM-Ing), ARTEMIS (IFEST). L'équipe étudie des techniques de modélisation et de validation formelle de systèmes.
L'équipe a intégré en janvier 2012 le laboratoire Lab-STICC (UMR-CNRS 6285). En particulier, l'ENSTA Bretagne contribue à apporter une expertise auprès du pôle Systèmes de Systèmes (SdS) lors des évaluations de projets (PEA KIMONO, OMOTESC) ou des métiers AESS, ISE. L'équipe participe au GT_NAF (Groupe de travail NAF).

L'équipe participe à des projets collaboratifs financés par l'ANR, la DGE et la communauté européenne dans le domaine de l'IDM (DOMINO, MOPCOM-SoC/SoPC, MOPCOM-Ing,

TOPCASED, ITEA-IFest). Ces projets sont menés en collaboration avec des partenaires académiques (INRIA, IRIT, ONERA, Supelec, Institut-Telecom, CEA-List) du domaine de l'IDM et des industriels préoccupés par l'intégration de ces techniques dans les développements de systèmes (THALES, AIRBUS, CNES, NEXTER, Thomson, Orange Labs, SODIUS, CS-SI).
L'équipe participe également à des projets dans le domaine FPGA (GDR Soc-Sip, ANR ARDyT) et développe des outils dans le domaine de l'embarqué et des FPGAs, avec parmi les applications visées, l'introduction d'accélérateurs de type FPGAs dans le Cloud Computing (Outil MADEO, IRT B-Com).

8.4 Contexte général

La cyber sécurité est un enjeu critique et sociétal important tant militaire que civil. De nombreux travaux ont été initiés sur le volet matériel (cryptographie, résilience, etc.) et sur le volet logiciel (bonnes pratiques de programmation, validation formelle, etc.).

Des méthodologies de capitalisation des connaissances sur les systèmes à sécuriser ou à découvrir, notamment sur l'interfaçage entre matériel et logiciel, restent à identifier et prototyper. Ces méthodologies doivent s'appuyer conceptuellement sur les cycles de conception traditionnels et bien maitrisés pour être transposer vers la sécurisation de systèmes connus et aussi à découvrir. Cet objectif doit permettre de réutiliser les connaissances antérieures et les adapter pour couvrir les enjeux de la cyber-sécurité.

8.5 Programme détaillé des travaux – planning et échéancier

La présente proposition se décline en 2 grands axes de travail :

- Le **premier axe de travail** se porte sur la partie matérielle de la proposition :
  - Le premier volet de cet axe vise à identifier les méthodes de conception classiques permettant de traiter la modélisation de l'articulation du matériel et du logiciel.
  - Un second volet vise à mettre en place un flot de synthèse système sur FPGAs pour permettre de prototypage d'architectures candidates.
- Le **second axe de travail** se porte sur la partie modélisation logicielle
  - Le premier volet de cet axe vise à identifier les langages de modélisation pour les différents niveaux d'abstraction du système. Ainsi que les langages et frameworks basés sur les rôles nécessaires à la fédération des modèles.
  - Le second volet vise à proposer une architecture de modèles fédérés afin de capitaliser la connaissance sur l'identification des interfaces matériel-logiciel.

Les deux axes seront déclenchés en parallèles. La demande de budget intègre donc le financement de deux ingénieurs/post-doctorants, porteurs de profils différenciés.

8.6 production scientifique relative au sujet (sur les 2 dernières années)

6. A Prototyping Platform for Virtual Reconfigurable Units. Lagadec L., Le Lann Jean-Christophe, Bollengier T. Recosoc 2014 - May, Montpellier France.
7. An experimental toolchain based on high-level dataflow models of computation for heterogeneous MPSoC. Julien Heulot, Karol Desnos, Jean-François Nezan, Maxime Pelcat, Mickaël Raulet, Hervé Yviquel, P.-L. Lagalaye, J-C Le Lann. DASIP'12
8. From system-level models to heterogeneous embedded systems, Jean-Christophe Le Lann, Joël Champeau, Papa Issa Diallo, Pierre-Laurent Lagalaye.RITF 2012 - Recherche et Innovation pour les Transports du Futur, Paris : France.
9. Modélisation algorithmique et synthèse d'architectures assistées par model-checking, Jean-Christophe Le Lann, Philippe Dhaussy, Pierre-Laurent Lagalaye.CAL 2012-, Montpellier : France.
10. MoPCoM Methodology: Focus on Models of Computation. Ali Koudri, Joël Champeau, Jean-Christophe Le Lann and Vincent Leilde. ECMFA'2010, Paris
11. Model federation in toolchains. J. Champeau, V. Leildé, and P. I. Diallo, in Workshop “Semantic Information Modeling for Federation “ in conjunction with MODELS 2013.
12. A Role Language to Interpret Multi-Formalisms System of Systems Models. Jean-Philippe Schneider, Joël Champeau, Ciprian Teodorov, Eric Senn and Loïc Lagadec. IEEE International System Conference, Vancouver, April 13-15, 2015.
13. Role Framework to Support Collaborative Virtual Prototyping of System of Systems. Jean-Philippe Schneider, Joël Champeau, Loïc Lagadec and Eric Senn. WETICE Conference June 15-17 2015.

9 - Partie financière 2015-2016	
Financement de l'opération (k€ TTC)	
Montant du projet 2015-2016 :	**146,8 KEuros** (ajout des coûts liés aux E/C permanents impliqués)
Soutien demandé à la DGA en 2015-2016 :	**112,8 KEuros**
Durée du projet (en mois) :	**12 mois (sur 36 mois au total)**

PRESENTATION DES DEPENSES
Préliminaire : le tableau ci-dessous identifie les montants demandés en financement de la DGA en regard du coût complet de l'opération pour l'année 2015-2016

1) *Dépenses de personnel*

Catégorie	Nombre d'hommes .mois	Coût unitaire mensuel	**Montant k€ HT**	**Montant k€ TTC**	**Montant financé par la DGA k€ TTC**
Post-docs (12 mois)	12	4,2 K€		50,4	50,4
Post-docs (12 mois)	12	4,2 K€		50,4	50,4
Total Dépenses de personnel					100,8

2) *Matériels consommables*
Indiquer la désignation des matériels

-		0	**0**
Total Dépenses de matériels consommables :			

3) *Matériels non consommables* .

-		0	**0**
Total Dépenses de matériels non consommables :			

4) *Frais de déplacements*

- Conférences			
Total Frais de déplacements			**10 K€**

5) *Autres Frais*
Préciser leurs justifications

Inscriptions conférences		0	**2 K€**
Total Autres Frais		0	**2 K€**