Simulation hybride de la sécurité des systèmes d'information*

Vers un environnement virtuel de formation

Guillaume Prigent, Fabrice Harrouet, Jacques Tisseau, and Frédéric Paul

Centre Européen de Réalité Virtuelle
Ecole Nationale d'Ingénieurs de Brest
25 rue Claude Chappe
BP 38, F-29280 Plouzané, France
{guillaume.prigent, harrouet, tisseau, frederic.paul}@enib.fr

1 Introduction

> *The lesson I learned was that security models and formal methods do not establish security. They only establish security with respect to a model, which by its very nature is extremely simplistic compared to the system that is to be deployed, and is based on assumptions that represent current thinking. Even if a system is secure according to a model, the most common (and successful) attacks are the ones that violate the model's assumptions. Over the years, I have seen system after system defeated by people who thought of something new.*
>
> Denning D.E., *The Limits of Formal Security Models*, [Denning 99]

Les systèmes d'information sont des systèmes complexes. Cette complexité provient essentiellement de la diversité des natures, de la diversité des connexions et de la diversité des interactions mises en jeu.

Un système d'information est alors *a priori* un milieu ouvert (apparition, disparition dynamique de composants, de nœuds, de connexions), hétérogène (morphologies et comportements variés) et formé d'entités composites, mobiles et distribuées dans l'espace comme dans le temps.

De part cette complexité, les systèmes d'information sont de plus en plus exposés aux différentes attaques et malversations des entités en interaction. Ces interactions, parmi lesquelles celles de l'homme avec son libre arbitre, peuvent être de natures différentes et opérer à différentes échelles spatiales et temporelles.

Les atteintes à la sécurité de tels systèmes peuvent être aussi bien interne au périmètre considéré, comme dans le cas des utilisateurs légitimes ou externes comme dans le cas des vagues d'attaques ou de sondages des vers chaque jour plus virulents. Gérer la sécurité de tels systèmes complexes est toujours coûteux en ressources, bien souvent éphémère, et parfois impossible.

Certes, la sécurité est une problématique entière des systèmes d'information mais, à notre sens, et basé sur notre propre expérience auprès d'élèves ingénieurs, elle ne peut être abordée que conjointement à une véritable prise de recul concernant le système d'information lui-même. En d'autres termes, la sécurité ne peut

* La plupart des travaux que nous présentons par la suite ont été réalisés dans le cadre du marché « EREBOR », entre la DGA/CELAR et le Centre Européen de Réalité Virtuelle

s'envisager que lorsque le socle des compétences requises, et en particulier de la connaissance des architectures des systèmes d'information, est déjà en place.

Paradoxalement, l'informatique propose que très peu d'outils permettant d'aider à appréhender la complexité des systèmes d'information et en particulier leur sécurité. Ceci peu s'expliquer, entre autre, parce qu'il est extrêmement difficile de formaliser ce type de système [Denning 99]. En effet, il n'existe pas aujourd'hui de théorie capable de formaliser cette complexité et de fait, il n'existe pas de méthodes de preuves formelles *a priori* comme il en existe dans les modèles hautement formalisés. En l'absence de preuves formelles, nous devons recourir à l'expérimentation du système d'information en cours d'évolution afin de pouvoir effectuer des validations expérimentales *a posteriori*.

En ce qui concerne la sécurité des systèmes d'information, un environnement virtuel de prototypage interactif, de formation, et de simulation des systèmes d'information serait une aide précieuse dans ce domaine.

De ce fait la composante formation d'un tel environnement nous apparaît à ce jour comme un des aspects les plus importants de nos objectifs. Pourtant, force est de constater que le domaine n'est que partiellement abordé. Soit la simulation et les environnements virtuels de formation concernent d'autres problématiques spécifiques à leurs cadres d'emplois, soit les outils du domaine de la sécurité proposent trop peu d'interactions et une abstraction insuffisante ou un périmètre restreint, incompatible avec nos objectifs.

Afin de tenter d'outiller de la sorte le domaine de la sécurité des systèmes d'information, il convient d'exposer l'historique de nos travaux afin de dégager la problématique de notre approche. Nous présentons notre plate-forme hybride de simulation des système d'information avant de proposer des perspectives qui nous semblent intéressantes.

2 Historique

> *Our hypothesis is that simulation-based computer security awareness training can be more focused and less expensive than lecture-lab courses in use today. Replacing the customized laboratory networks with a simulation can address the cost factor. Simulation based training can be more effective because it allows several things that are impossible with today's lecture-lab training.*
>
> Tanner M.C., Elsaesser C., Whittaker G.M., *Security Awareness Training Simulation*, [Tanner et al 01]

Depuis début 1999, notre équipe de recherche a abordé la problématique de la simulation de la sécurité des systèmes d'information suivant différents axes. Ces différents travaux ont pour la plupart été initiés par l'équipe SSI du Centre Electronique de l'Armement. La partie la plus significative et l'état actuel de notre réflexion concerne le projet EREBOR qui s'est déroulé sur une période de 24 mois.

Nous avons volontairement effectué nos travaux exploratoires avec d'autres approches que celles habituelles dans le domaine de la recherche en sécurité des systèmes d'information. En voici les principaux projets :

- G-Fox 1 : Simulation Multi-Agents de la disponibilité des systèmes d'information,

– G-Fox 2 : Simulation Multi-Agents de la confidentialité et de l'intégrité des systèmes d'information,
– EREBOR : Maquette d'environnement virtuel de simulation hybride de la sécurité des systèmes d'information.

Les conclusions du projet EREBOR, nous ont encouragé dans la voie du couplage entre le réel et le virtuel au sein d'une simulation. C'est à partir de cette réflexion que nous avons restreint et formalisé plus précisément notre problématique.

3 Problématique

> *The high cost of running real-world attacks, the limited extent to which they exercise the space of actual attacks, and the high potential for harm from a successful attack conspire to make some other means of analysis an imperative.*
> Cohen F., *Simulating Cyber Attacks, Defenses, and Consequences*, [Cohen 99]

La réalité virtuelle fournit aujourd'hui un cadre conceptuel, méthodologique et expérimental bien adapté pour imaginer, modéliser et expérimenter la complexité des systèmes d'information.

Associé à un couplage hybride fort entre le réel et le virtuel au niveau des éléments constitutifs d'un système d'information, la réalité virtuelle permet d'ajouter une composante d'immersion dans les modèles, qui nous apparaît bénéfique pour une meilleure prise en compte de la complexité des systèmes d'information.

L'ambition de nos objectifs impose la mise en évidence des principaux écueils techniques relatifs au couplage réel/virtuel.

3.1 Réalité virtuelle et expérimentation

> *La réalité virtuelle est un domaine scientifique et technique exploitant l'informatique et des interfaces comportementales en vue de simuler dans un monde virtuel le comportement d'entités 3D, qui sont en interaction en temps réel entre elles et avec un ou des utilisateurs en immersion pseudo-naturelle par l'intermédiaire de canaux sensori-moteurs.*
> Fuchs P., Arnaldi B., Tisseau J., *La réalité virtuelle et ses applications*, [Fuchs et al. 03]

Selon la définition précédente, la réalité virtuelle place l'utilisateur au cœur d'un véritable laboratoire virtuel qui le rapproche ainsi des méthodes des sciences expérimentales tout en lui donnant accès aux méthodes numériques.

Dépassant la simple observation de l'activité du modèle numérique en cours d'exécution sur un ordinateur, l'utilisateur peut tester la réactivité et l'adaptabilité du modèle en fonctionnement, tirant ainsi profit du caractère comportemental des modèles numériques. Nous appelons ce nouveau type d'expérimentation : l'expérimentation *in virtuo*.

Une expérimentation *in virtuo* est ainsi une expérimentation conduite dans un univers virtuel de modèles numériques en interaction et auquel l'homme participe. La réalité virtuelle implique pleinement l'utilisateur dans la simulation, rejoignant ainsi l'approche de la conception participative (*participory design* [Schuler et al. 93]) qui préfère voir dans les utilisateurs des *acteurs humains* plutôt que des *facteurs humains* [Bannon 91]. Une telle simulation participative

en réalité virtuelle met en œuvre des modèles de types différents (multi-modèles) issus de domaines d'expertise différents (multi-disciplines).

Elle est souvent complexe car son comportement global dépend autant du comportement des modèles eux-mêmes que des interactions entre modèles. Enfin, elle doit inclure le libre arbitre de l'utilisateur humain qui exploite les modèles en ligne.

L'expérimentation *in virtuo* implique ainsi un vécu que ne suggère pas la simple analyse de résultats numériques. Entre les preuves formelles *a priori* et les validations *a posteriori*, il y a aujourd'hui la place pour une réalité virtuelle vécue par l'utilisateur qui peut ainsi franchir le cap des idées reçues pour accéder à celui des idées vécues.

3.2 Couplage Hybride Réel/Virtuel

> *The high cost of running real-world attacks, the limited extent to which they exercise the space of actual attacks, and the high potential for harm from a successful attack conspire to make some other means of analysis an imperative.*
> Cohen F., *Simulating Cyber Attacks, Defenses, and Consequences*, [Cohen 99]

Dans le domaine des systèmes d'information, et en particulier celui des réseaux informatiques, le terme « simulation hybride » évoque généralement un procédé de simulation de tels réseaux qui utilise des méthodes de calculs empruntées à la physique reposant à la fois sur des modèles continus (de type écoulement de fluides) et des modèles événementiels [Bell et al. 78]. Cette approche est principalement axée sur le dimensionnement des moyens de communication et repose sur des données quantitatives telles le débit ou le temps de réponse [Kiddle et al. 03].

Pour les travaux que nous envisageons, le terme hybride prend sa signification dans le contexte d'un couplage entre des éléments de systèmes d'information réels et simulés. Il s'agit ainsi de se donner les moyens d'expérimenter sur un système d'information qui soit matérialisé par l'interconnexion de matériel et logiciel existant avec du matériel et du logiciel simulé. Ceci peut être exploité selon trois types de configurations :

- simuler un système d'information complet,
- faire intervenir des élements simulés dans une architecture existante,
- faire intervenir des élements réels dans une architecture simulée.

Les opérations menées sur les éléments réels doivent ainsi subir l'influence des éléments simulés (comme s'ils étaient réels). Les interventions des opérateurs humains sur les éléments simulés doivent mettre en œuvre les concepts et les outils usuellement employés en situation réelle.

Nous pouvons alors parler de « réalité mixte » dans le sens ou les interventions des opérateurs humains prennent place à la fois dans un contexte réel et simulé permettant ainsi de raisonner et réagir dans des situations perçues comme réelles sur des infrastructures partiellement ou pas du tout réalisées.

3.3 Objectifs

> *For airline pilots simulators provide scenarios that would be much too risky to duplicate in the "real" world. The FAA gives equal credit to pilots for time spent in a simulator*

as time flying a real aircraft. Simulations of the interaction of factors in large scale, long-term projects have yielded tens of millions of dollars in savings.
Saunders J.H.,*« The Case for Modeling and Simulation of Information Security »*,[Saunders 01]

Un système complexe a un comportement singulier : il est impossible de restituer de manière reproductible l'état et le comportement d'un tel système. Ce qui lui confère un statut différent des systèmes étudiés selon une approche réductionniste classique qui nécessite un déterminisme fort, est donc la reproductibilité du phénomène. La modélisation d'un système complexe produit alors un modèle qui vient se substituer au système lui-même, et la simulation de ce modèle doit permettre d'anticiper au mieux le comportement du système réel. Ne disposant pas de modèles formels prédictifs pour un système complexe, la simulation repose donc sur l'existence de modèles numériques et/ou analogiques, et/ou sur la possibilité d'expérimenter le système lui-même.

C'est le cas des systèmes d'information pour lesquels on peut disposer du système lui-même, de modèles numériques (maquettes virtuelles) et de modèles analogiques (maquettes réelles). Il est donc envisageable de fusionner ces différentes possibilités pour une meilleure compréhension et une meilleure analyse de ces systèmes. Plusieurs architectures de couplage hybride sont envisageables. Elles reposeront toutes sur une simulation à base d'itérations asynchrones et chaotiques et se heurteront toutes à la capacité d'interconnexion des systèmes d'information virtuels et réels tant du point de vue matériel que logiciel.

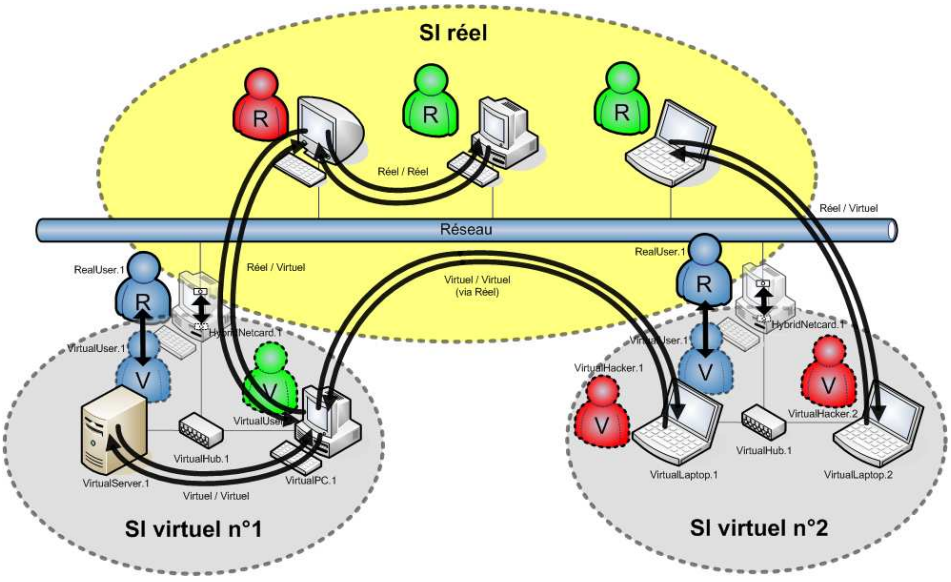

Fig. 1. Objectif général d'interconnexion réel/virtuel

Une telle démarche permet d'envisager un outil donnant une vision globale d'un système d'information en devenir. L'objectif principal consiste à placer les opérateurs humains en situation dans un tel système afin qu'il en « ressente » le fonctionnement bien avant qu'il soit pleinement réalisé.

3.4 Difficultés techniques du couplage

La réalisation technique d'un tel couplage entre l'infrastructure réelle et l'infrastructure simulée repose principalement sur deux points :

- être capable de recevoir et de produire, sur le système hôte hébergeant la simulation, des trames réseau qui ne concernent pas l'hôte en lui même mais les éléments simulés,
- pouvoir réutiliser, sur le système hôte, des logiciels réseau existants de telle sorte qu'ils utilisent les propriétés réseau des dispositifs simulés en lieu et place de celles du système hôte.

Le premier point technique, pouvant être qualifié de « couplage matériel », est relativement bien maîtrisé ; il suffit d'utiliser des *sockets* de niveau liaison (*ethernet*) en mode *promiscuous*. Ceci permet de capturer toutes les informations qui parviennent aux interfaces réseau physiques du système hôte et d'émettre sur ces interfaces des trames fabriquées à notre convenance.

Le couplage matériel permettant de recevoir et d'émettre sur la machine hôte des trames qui ne la concernent pas directement étant techniquement possible, il nous est désormais nécessaire de nous interroger sur la manière dont les machines simulées doivent consommer et produire de telles trames afin de présenter sur le réseau des comportements similaires à ceux que produiraient des machines réelles.

Ces comportements sont essentiellement la manifestation du fonctionnement d'outils (du domaine des réseaux) sur de telles machines. Dans la pratique, ces outils sont principalement des clients et des serveurs (*HTTP*, *SSH*, ...) ou des outils d'investigation voire d'attaque (*Nmap*[1], *Nessus*[2], ...). Etant donné que ces machines sont simulées par un programme informatique, la première démarche qui pourrait sembler envisageable consisterait à simuler également l'exécution de ces outils.

Bien que techniquement possible, cette solution représente un travail de réécriture colossal qui est raisonnablement irréalisable dans la pratique. Une solution qui permettrait de réutiliser les outils existants semble plus intéressante pour deux raisons. Tout d'abord cela éviterait le travail de réécriture évoqué ici. Le second intérêt de cette solution vient du fait que les outils utilisés en l'état conserveront leurs caractéristiques propres (notamment leurs vulnérabilités).

Ce dernier point va tout à fait dans le sens d'une simulation qui soit aussi semblable que possible à la réalité.

[1] « Nmap - Free Security Scanner For Network Exploration », http ://www.insecure.org/nmap/

[2] « Nessus - Open Source Security Scanner », http ://www.nessus.org/

Toutefois, l'utilisation pure et simple des outils existants implique leur exécution dans le contexte de la machine hôte qui normalement doit être complètement étrangère aux machines simulées. Ceci implique entre autres que les accès au réseau se feront par l'intermédiaire du système et des ressources de la machine hôte, c'est à dire que les trames échangées concerneront l'identité de la machine hôte, les protocoles de résolution d'adresse (*ARP*) et de nom (*DNS*) et le routage seront également dépendant de la configuration et de l'état de la machine hôte.

Il est donc nécessaire de trouver un moyen pour que l'exécution de ces outils n'interfère pas avec le système hôte et à ce propos, il est légitime de s'interroger sur la possibilité d'avoir recours à des émulateurs complets tels que *VMware*[3], *Bochs*[4], *PearPC*[5], *UML*[6] ... Ceux-ci permettent effectivement d'émuler une machine complète dans laquelle il est possible de faire fonctionner des systèmes variés.

Cependant l'utilisation de plusieurs de ces émulateurs sur la même machine hôte est extrêmement pénalisante en terme de performances. De plus, chaque machine émulée doit être configurée et utilisée une à une comme s'il s'agissait d'autant de machines physiques. Il n'est pas question dans ces conditions d'avoir un logiciel de simulation unifié qui représente l'ensemble d'un parc informatique (machines, commutateurs, ponts, connectique, ...) reconfigurable à volonté. On fait juste l'économie du prix (les licences logicielles devant toujours être en règle) et de l'encombrement des machines physiques. Nous ne retiendrons donc pas cette solution d'émulation de système complet et nous nous concentrerons uniquement sur la simulation de la partie réseau qui constitue le centre d'intérêt principal de nos travaux.

Il est possible de créer sur le système hôte des interfaces réseau virtuelles qui ne sont directement associées à aucun périphérique de communication. C'est le cas par exemple des réseaux privés virtuels (*VPN*) qui traitent les trames reçues et émises sur une interface virtuelle pour les insérer dans un flux véhiculé par une interface physique. De telles interfaces virtuelles peuvent être configurées et ainsi se voir attribuer leurs propres adresses matérielle (*MAC*) et réseau (*IP*). Il serait donc tentant d'attribuer à chaque machine simulée une telle interface afin que les outils exécutés communiquent via celle-ci. Seulement, cette solution n'est envisageable que pour un nombre très restreint d'outils (principalement d'investigation et d'attaque) qui permettent de choisir explicitement l'interface réseau à utiliser. La plupart des applications clientes ou serveurs utilisent l'API socket pour une communication de niveau transport (*TCP* ou *UDP*) ; or à ce niveau d'abstraction, le choix de l'interface réseau à utiliser est entièrement délégué au système. De plus, les résolutions *ARP* et *DNS* ainsi que le routage sont toujours entièrement pris en charge par le système hôte (les interfaces virtuelles jouent le

[3] « VMware - Virtual Infrastructure Software », http ://www.vmware.com/
[4] « bochs - The Open Source IA-32 Emulation Project », http ://bochs.sourceforge.net/
[5] « PearPC - PowerPC Architecture Emulator », http ://pearpc.sourceforge.net/
[6] « UML - User Mode Linux », http ://usermodelinux.org/

même rôle que les interfaces physiques). Cette solution ne pourra donc pas être retenue.

Nous venons ici d'identifier le point crucial de notre problème de couplage logiciel : la pile *TCP/IP* du système hôte forme une entité unique qui n'est exploitable de manière cohérente que par une machine unique. Si nous souhaitons conserver les services réseau en bon état de fonctionnement pour la machine hôte, il est nécessaire de fournir une nouvelle pile *TCP/IP* pour chaque machine simulée. L'implémentation de celle-ci doit être totalement décorellée de celle du système hôte afin qu'il n'y ait aucune interférence.

La seule interaction doit se situer au niveau de la capture et de l'injection de trames par le couplage matériel déjà présenté. L'implémentation d'une telle pile *TCP/IP* représente un travail délicat mais tout à fait envisageable si l'on se concentre sur les fonctionnalités essentielles. Il reste donc à discuter la manière dont on peut contraindre les outils à utiliser cette pile *TCP/IP* en lieu et place de celle du système hôte.

Dans un premier temps il paraît raisonnable de recenser les outils dont nous souhaitons disposer dans le cadre de l'expérimentation autour de la sécurité des systèmes d'informations. Ainsi ces quelques outils pourront être très partiellement modifiés (dans leur code source) afin que les appels à l'*API socket* usuels soient remplacés par des appels à la nouvelle implémentation. Le nombre de ces modifications est dans la pratique très restreint ; il ne s'agit la plupart du temps que des opérations d'ouverture, de fermeture, de lecture, d'écriture et d'attente sur les canaux de communications.

Un raffinement supplémentaire pourrait consister à détourner automatiquement ces appels par le mécanisme de préchargement de symboles liés dynamiquement. Cette dernière solution offre l'avantage d'éviter d'avoir à adapter les outils mais elle nécessite alors que l'implémentation de la pile *TCP/IP* simulée soit en mesure de gérer toutes les subtilités de l'*API socket* (*ioctl*, *fcntl*, *setsockopt*, ...) auxquelles les applications pourraient avoir recours. Ce dernier point représente un volume de travail comparable à la réécriture d'une portion non négligeable d'un système d'exploitation. Ce raffinement ne pourrait donc être envisagé dans la pratique qu'après avoir modifié un grand nombre d'outils afin de recenser petit à petit parmi toutes ces subtilités lesquelles sont les plus souvent utilisées.

Une difficulté technique supplémentaire concerne la cohérence temporelle de la simulation. Tout d'abord, les éléments simulés doivent interagir dans une échelle de temps qui reste compatible avec les éléments réels afin que les opérateurs humains puissent se sentir en situation opérationnelle ; dans ces conditions la simulation a lieu en « temps réel ». Toutefois, un autre mode de simulation doit pouvoir être envisagé ; il s'agit d'un « temps virtuel » dans lequel chaque opération doit être datée selon un temps qui n'a rien à voir avec le temps de calcul nécessaire à la simulation de cette opération. Dans ces conditions, il est alors envisageable de quantifier très finement les caractéristiques des flux en terme de débit, de temps de réponse ... mais ceci ne peut se faire que dans la cadre de la simulation sans aucun couplage avec les dispositifs physiques. Cette

seconde difficulté est d'un ordre plus théorique que la précédente et nécessite principalement un travail de recherche.

4 L'environnement virtuel BridNet

> *On a different level though, and perhaps as a base, security professionals would benefit through a larger understanding of basic concepts in modeling such as levels of abstraction, logical versus physical entities, objects attributes, and scripting.*
> Saunders J.H., *The Case for Modeling and Simulation of Information Security*, [Saunders 01]

La plate-forme BridNet[7] a été conçue dans le but de permettre l'étude de la sécurité des systèmes complexes que sont les systèmes d'information. Se basant sur les concepts de simulation et de simulation hybride au sens du couplage entre le réel et le virtuel, ses objectifs sont de fournir un environnement virtuel d'expérimentation adapté aux futurs professionnels œuvrant dans le domaine des systèmes d'information et de leur sécurité.

Cet outil repose sur le concept d'étude de la complexité dans un laboratoire virtuel et par l'expérimentation *in virtuo*. La simulation des systèmes d'information ainsi considérée permet à l'utilisateur spectateur-acteur-créateur d'observer, d'interrompre ou de perturber le système étudié, mais aussi de tester l'adaptabilité du modèle en fonctionnement et son interopérabilité avec les systèmes environnants.

L'environnement BridNet permet la conception de systèmes d'information virtuels, et leur interaction avec d'autres systèmes d'information réels ou virtuels de manière transparente. En ce sens cet outil peut être vu comme un "honeynet" ([Spitzner 02][Spitzner 03][Stella et al. 04]) d'un point de vue du système réel en interaction avec lui.

4.1 Mise en œuvre technique

Afin de répondre aux objectifs d'expérimentation *in virtuo*, d'immersion et d'interaction de l'apprenant dans et avec le modèle simulé, il a été développé diverses représentations de ce modèle. C'est ainsi que différentes IHM et vues du modèle ont été créées, dont une en 2D facilitant l'interaction et la visualisation de l'organisation du modèle, et une en 3D permettant d'appréhender la représentation physique des éléments du système, par immersion dans le modèle simulé.

L'interaction avec les modèles réalisée, il restait à mettre en œuvre le couplage hybride réel/virtuel. Cette partie assez délicate, puisque cœur du système, a nécessité l'écriture d'une pile réseau simplifiée. Il aurait en effet été raisonnablement impossible et inutile de réécrire une pile complète. L'usage du mode promiscuous permettant le transfert des flux du système réel à cette pile réseau (reliée logiquement aux autres entités du modèle simulé), et inversement, permet dès lors le couplage réel/virtuel attendu. Une pile réseau virtuelle est alors disponible à chaque élément du modèle simulé qui doit être relié au réseau réel.

[7] « BridNet - hybBrid Network »

Les difficultés techniques et les fonctionnalités désirées liées à la gestion du temps au niveau du déroulement de la simulation ont donné naissance à deux modes de fonctionnement temporel : les modes « temps réel » et « temps virtuel ». Deux ordonnanceurs ont donc été implémentés, l'un permettant un fonctionnement « temps réel » *best-effort*, dont l'objectif est d'assurer au mieux la cohérence temporelle d'un modèle simulé avec un système réel, l'autre permettant une dilatation du temps et ainsi l'observation des événements liés principalement au modèle.

4.2 Fonctionnalités

La mise en œuvre technique permet déjà d'avoir un aperçu des fonctionnalités de la plate-forme : interaction, immersion, couplage hybride réel/virtuel, fonctionnement « temps réel »/« temps virtuel ».

Afin de rendre plus crédible et fonctionnel le réseau du point de vue extérieur, des mécanismes ont été intégrés pour gérer la disponibilité de services, simulés sur chaque entité du modèle. Il est alors possible de rediriger des flux réseau vers des services virtuels, ou encore, tout comme sur honeyd, de rediriger des flux vers des scripts ou applications système.

La plate-forme étant également orientée formation et étude de systèmes, d'autres mécanismes et fonctionnalités ont été implémentés afin de d'aider à la compréhension du modèle en fonctionnement. Un système de détection d'intrusions basé sur Snort a été mis en place, permettant de générer et de visualiser les alertes en cas d'attaque circulant dans le modèle simulé. De plus, chaque entité du modèle a la possibilité d'enregistrer les flux réseau reçus pour analyse ultérieure. Ces flux sont alors enregistrés au format "pcap", ce qui permet l'utilisation d'outils externes tels que tcpdump ou ethereal pour étudier à postériori les flux ayant circulé dans le réseau virtuel.

Diverses autres fonctionnalités purement applicatives ont également été développées, visant à simplifier l'utilisation de l'application, ou son administration. Par exemple, les mécanismes de sérialisation/désérialisation qui permettent l'enregistrement/chargement de systèmes virtuels, ainsi que de modèles de systèmes (appelés « templates ») pré-enregistrés par les utilisateurs ; ou encore le gestionnaire d'authentification centralisé, qui permet de restreindre l'accès à l'application ou à la base des modèles.

De nombreuses autres fonctionnalités et améliorations, telles que l'usage des outils existant depuis le modèle simulé, sont à l'état expérimental ou en prévision, mais ouvrent des perspectives prometteuses.

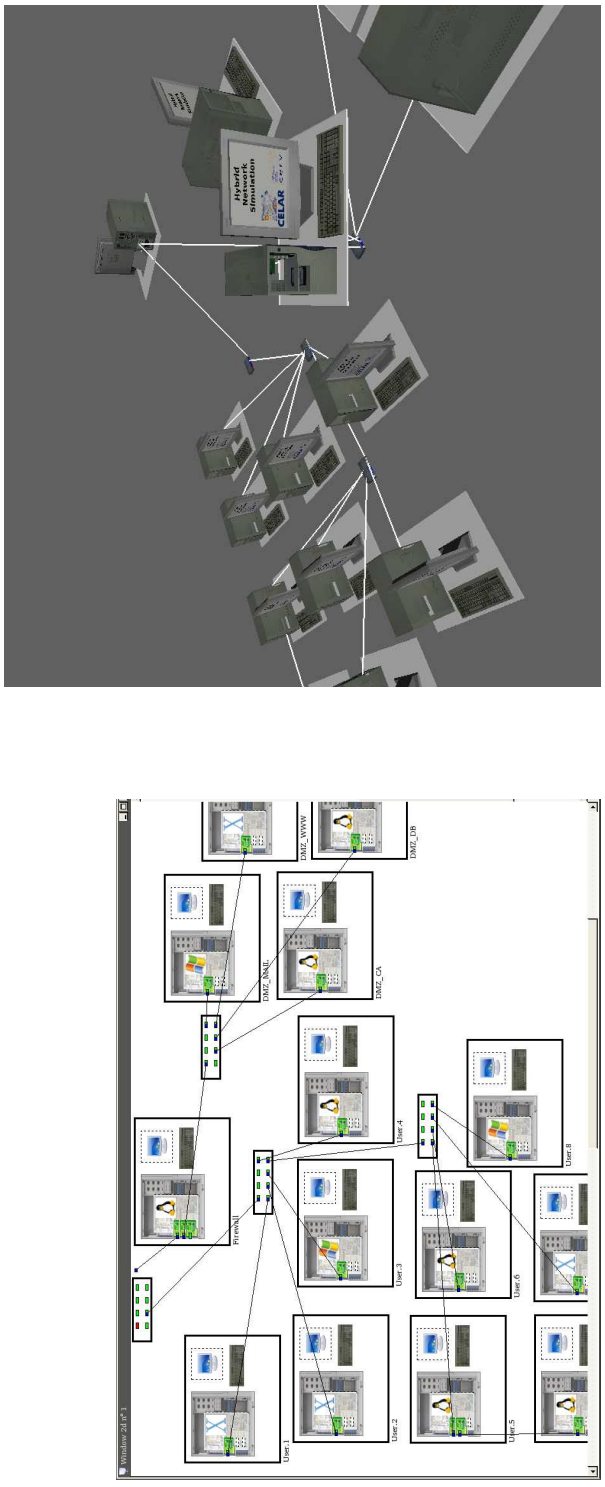

Fig. 2. Représentations 2D et 3D du système virtuel

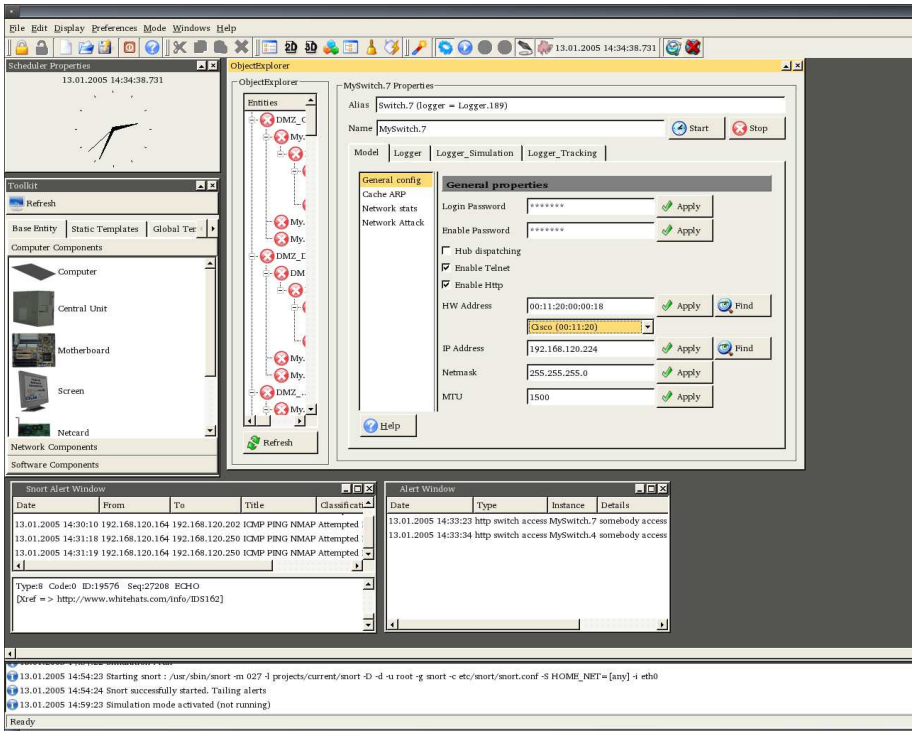

Fig. 3. Aperçu de l'IHM BridNet

5 Perspectives

> *It appears that these initial results are only the beginning of the sorts of results that simulation technology will provide in the information protection field, and that it is a fruitful area to explore.*
>
> Cohen F., *Simulating Cyber Attacks, Defenses, and Consequences*, [Cohen 99]

L'approche que nous avons exposée et illustrée au travers d'une maquette d'environnement virtuel permet d'aborder sous un nouvel angle le domaine de la formation à la conception et à la maintenance de systèmes d'information ; en effet, le déploiement d'un plateforme de tests complète nécessite des moyens techniques et financiers conséquents. Toutefois l'expérimentation en situation représente une étape importante dans le procédé de formation : on apprend en faisant ! [N'Guyen 98]

De nombreux domaines de formation autre que la sécurité des systèmes d'information, tels que la conduite automobile ou la formation des pompiers professionnels, nécessitent la mise en situation des apprenants ; ceux-ci doivent acquérir non seulement des connaissances, mais encore de véritables compétences. Cette compétence consistant à résoudre des problèmes en situation dynamique (incertaine, évolutive, et à forte contrainte temporelle) est particulièrement diffi-

cile à aborder par une formation classique : étude de cas, proposition de règles générales, instructions relatives à des scenarii probables...

Au contraire, la simulation informatique en général et la réalité virtuelle en particulier permettent d'immerger les apprenants dans les environnements où ils peuvent essayer, choisir, prendre des initiatives, échouer et recommencer. La confrontation aux situations pourrait leur permettre d'élaborer en mémoire à long terme des schémas d'actions articulant divers composants de la compétence : des connaissances générales liées au domaine, des stratégies ou des règles contextualisées, des procédures cognitives et des savoir-faire. La mise en œuvre d'élements simulés permet notamment le déroulement de scénarios préétablis ou pilotés par un formateur, afin de simuler des défaillances et ainsi mettre l'apprenant à l'épreuve. Ce dernier point trouve tout à fait sa justification dans la problématique de la formation à la sécurité des systèmes d'information [Saunders 02].

Au niveau de la virtualisation de processus réels fonctionnant sur la machine hôte de la simulation, la piste présentée qui nous semble la plus prometteuse concerne le détournement des appels systèmes réseau des processus réels. Nos premiers essais nous encouragent dans cette voie d'autant plus fortement que c'est exactement la même approche qui semble avoir été prise par l'équipe de *honeyd*[8] concernant la directive *subsystem*[9].

Au niveau de l'architecture d'un tel environnement et particulièrement dans le cas de la formation, il pourrait être intéressant de proposer une distribution de l'application. Dans cette optique il s'agirait de proposer un serveur de modèles auquel se connecteraient des clients de simulation. Cette architecture traditionnelle client/serveur permettrait ainsi d'envisager rapidement des applications distribuées de simulation hybride des systèmes d'information. À titre d'exemple et dans le domaine de la formation, ce type de modèle peut permettre la mise en place de sessions collaboratives d'enseignement assisté a distance (plusieurs clients partagent le même modèle, lui-même supervisé par un formateur). Bien sûr ce type de distribution ne peut être envisagé que dans une architecture « ad-hoc ». Le choix des messages à distribuer et des interactions à prendre en compte apparaît comme primordial. Il semble peu réaliste de vouloir distribuer la sémantique complète des paquets échangés (ce qui revient à échanger l'ensemble du traffic réseau de la plate-forme serveur aux clients). Disposer d'informations comme les bandes passantes occupées ou les attaques détectées peut suffire dans un cadre de formation ou d'architecture haut niveau d'un système d'information.

Comme pour la distribution des modèles, une perspective parmi d'autres concerne la distribution des interfaces monde réel/monde virtuel. Aujourd'hui une interface hybride comme nous la définissons est un périphérique réseau matériel en mode *promiscuous* afin d'utiliser des *sockets* de niveau liaison (*ethernet*). Une extension de ce couplage matériel à des périphériques distants (une forme de serveur de paquets) permettrait de pouvoir coupler des mondes virtuels locaux à la simulation à des systèmes d'information réel distants. La distribution des périphériques de couplage matériel pose les mêmes problèmes que dans le

[8] « Honeyd Virtual Honeypot », http ://www.honeyd.org/
[9] http ://www.honeyd.org/general.php

cas de la distribution des modèles : celui de l'occupation de la bande passante. La mise en place de tels canaux qui ne pertuberait pas l'environnement virtuel de simulation lui-même semble nécessiter la séparation physique des réseaux et des interfaces réseaux de la machine accueillant l'environnement virtuel.

En l'état actuel de nos travaux, nous cherchons principalement à donner une existence bien réelle dans un système d'information réel d'entités purement virtuelles. Une approche complémentaire pourrait être de créer dans l'environnement de simulation une représentation virtuelle des entités réelles des systèmes d'information réels couplés. Sur le même principe, il est envisageable d'utiliser des ressources réelles dédiées apparaissant comme virtuelles dans notre simulation. Cette dernière fonctionnalité s'apparente au niveau réseau à une translation de ports. Ce cas peut par exemple s'avérer utile lorsque la granularité nécessaire dépasse le réalisable en purement virtuel. Cette approche semble intéressante pour fournir une meilleure immersion à l'utilisateur de l'environnement virtuel.

Enfin, si nos objectifs étaient jusqu'alors de définir le contenant et créer un environnement virtuel dédié à l'architecture des systèmes d'information, il convient de s'intéresser aux contenus véhiculés dans notre environnement. L'humain avec son libre arbitre fait partie intégrante des composantes des systèmes d'information. De ce fait, des utilisateurs virtuels, avec leur part d'autonomie, doivent nécessairement être présents dans nos systèmes d'information virtuels. C'est principalement leurs comportement et leurs interactions avec les entités virtuelles (comme réelles) qui génère la « vie » et par là même le contenu des systèmes simulés ou couplés. À l'utilisateur réel de l'environnement virtuel est alors associé un avatar virtuel. Le Centre Européen de Réalité Virtuelle travaille depuis longtemps, cela au travers de différentes équipes et de différents projets, à l'autonomisation des modèles. C'est cette composante humaine et comportementale qui peut à notre sens réserver l'intérêt majeur pour la sécurité des systèmes d'information d'un environnement virtuel de simulation hybride comme nous l'entendons.

Au delà du cadre de la formation, un tel outil peut également servir à prototyper les infrastructures qui devront être déployées lors de missions particulières. L'interconnexion de systèmes existants peut également être abordée en associant un dispositif réel avec son pendant simulé. La démarche de simulation hybride peut donc être vue dans ce contexte comme un outil d'aide à la décision permettant de guider dans la mise en place d'architectures de systèmes d'information sécurisés.

Références

[Bannon 91] Bannon L.J., *From human factors to human actors : the role of psychology and human-computer interaction studies in system-design*, dans [**?**] :25-44, 1991

[Bell et al. 78] Bell G., Fuller S.H., Siewiorek D. (editors), *Hybrid Simulation Models of Computer Systems*, Communications of the ACM, volume 21, number 9, 1978

[Cohen 99] Cohen F., *Simulating Cyber Attacks, Defense, and Consequences*, http ://all.net/journal/ntb/simulate/simulate.html, 1999

[Denning 99] Denning D.E., *The Limits of Formal Security Models*, National Computer Systems Security Award Acceptance Speech, http ://www.cs.georgetown.edu/ denning/infosec/award.html, 1999

[Fuchs et al. 03] Fuchs Ph., Arnaldi B., Tisseau J., *La réalité virtuelle et ses applications*, dans *Le traité de la réalité virtuelle*, $2^{ème}$ édition, volume 1, chapitre 1, pages 3-52, Les Presses de l'Ecole des Mines de Paris, 2003

[Kiddle et al. 03] Kiddle C., Simmonds R., Williamson C., Unger B., *Hybrid Packet/Fluid Flow Network Simulation*, ACM 17th International Workshop on PADS, pages 143-152, San Diego, 2003

[Kiehl et al. 03] Kiehl T., Mattheyses R., Simmons M., *Hybrid Simulation of Cellular Behavior*, Rapport interne, Advanced Computing Technologies, Niskayuna, 2003

[N'Guyen 98] N'Guyen, *Les mécanismes d'apprentissage*, Revue Française de Pédagogie, page 112, 1998

[Saunders 01] Saunders J.H., *The Case for Modeling and Simulation of Information Security*, Computer Security Institute Conference, http ://www.johnsaunders.com/papers/securitysimulation.htm, 2001

[Saunders 02] Saunders J.H., *Simulation Approaches in Information Security Education*, Proceedings of the National Colloquium for Information Systems Security Education, http ://www.johnsaunders.com/papers/ncisse/ncisse2002paper.pdf, 2002

[Schuler et al. 93] Schuler D., Namioka A. (editors), *Participatory Design : Principles and Practices*, Lawrence Erlbaum Associates, Hillsdale, 1993

[Spitzner 02] Spitzner L., *Honeypots : Tracking Hackers*, Addison-Wesley, ISBN 0321108957, 2002

[Spitzner 03] Spitzner L., *Honeypots : Definitions and Value of Honeypots*, http ://www.tracking-hackers.com, 2003

[Stella et al. 04] Stella E., Martineau T., *Specter : un honeypot qui compromet les pirates ; techniques et légalité*, MISC n°11, janvier-février, pages 6-9, 2003

[Tanner et al 01] Tanner M.C., Elsaesser C., Whittaker G.M., *Security Awareness Training Simulation*, The MITRE Corporation, 2001