Intégration d'un dispositif d'apprentissage et de classification dans un système d'évaluation de la sécurité informatique en environnement de Cloud Computing

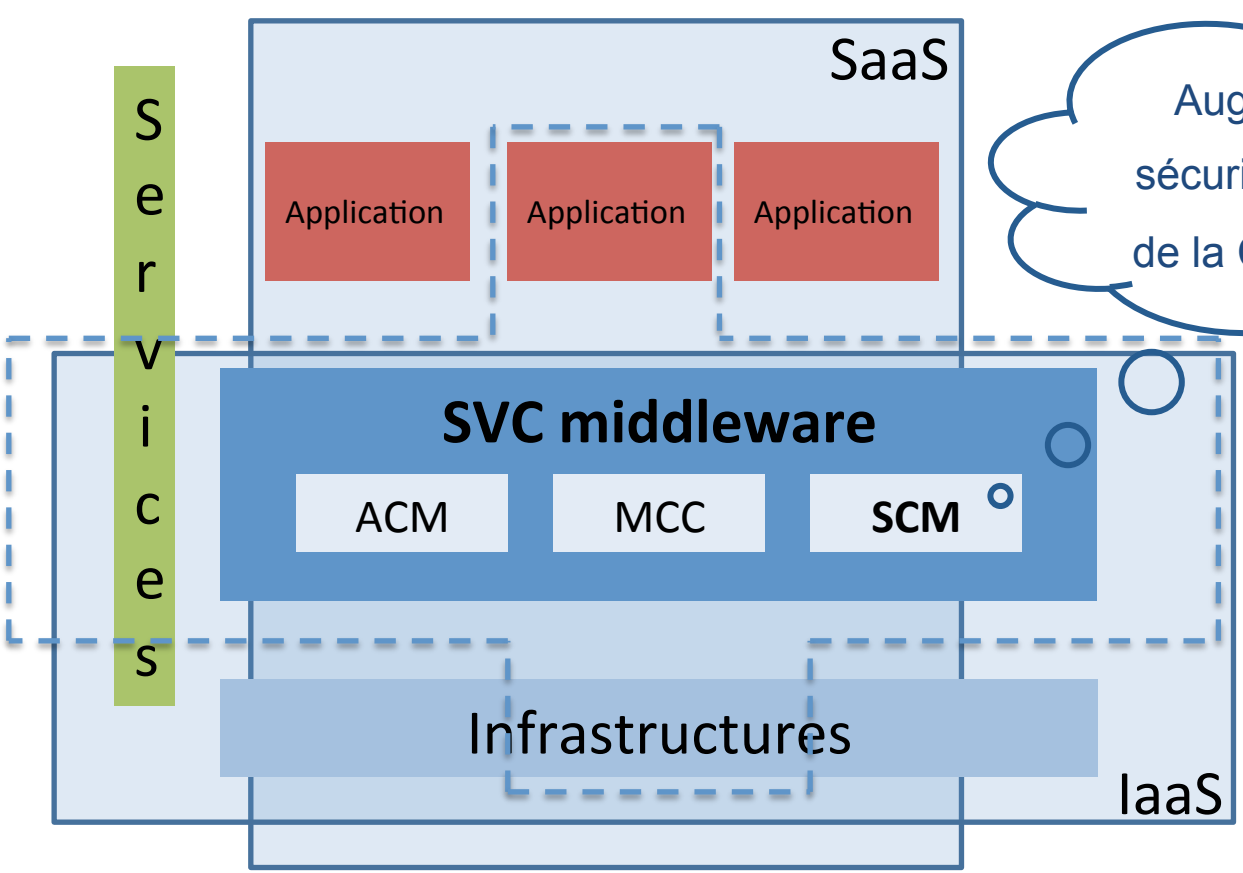
Meriam MAHJOUB

Équipe VORTEX, IRIT UMR 5505

16 Décembre 2013

Avancement de thèse

Plan



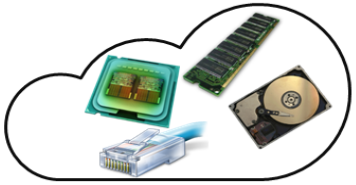
Augmenter le niveau de sécurité et de confidentialité de la Couche IaaS du Cloud

Périmètre de SVC

- ☺ A tout moment
- ☺ De n' importe où
- ☺ Payer selon la consommation

- ☺ Virtualisation
- ☺ Délégation de la gestion du matériel

Un modèle permettant d'accéder à travers un réseau et à la demande à un ensemble de ressources informatiques partagées.



Ressources partagées

Des ressources virtuelles affectés dynamiquement en fonction des besoins des clients.

Accès réseau

Des services accessible via le réseau, quelque soit le périphérique.

Services à la demande

Possibilité de consommer des services ou des ressources sans demander l'intervention du fournisseur.

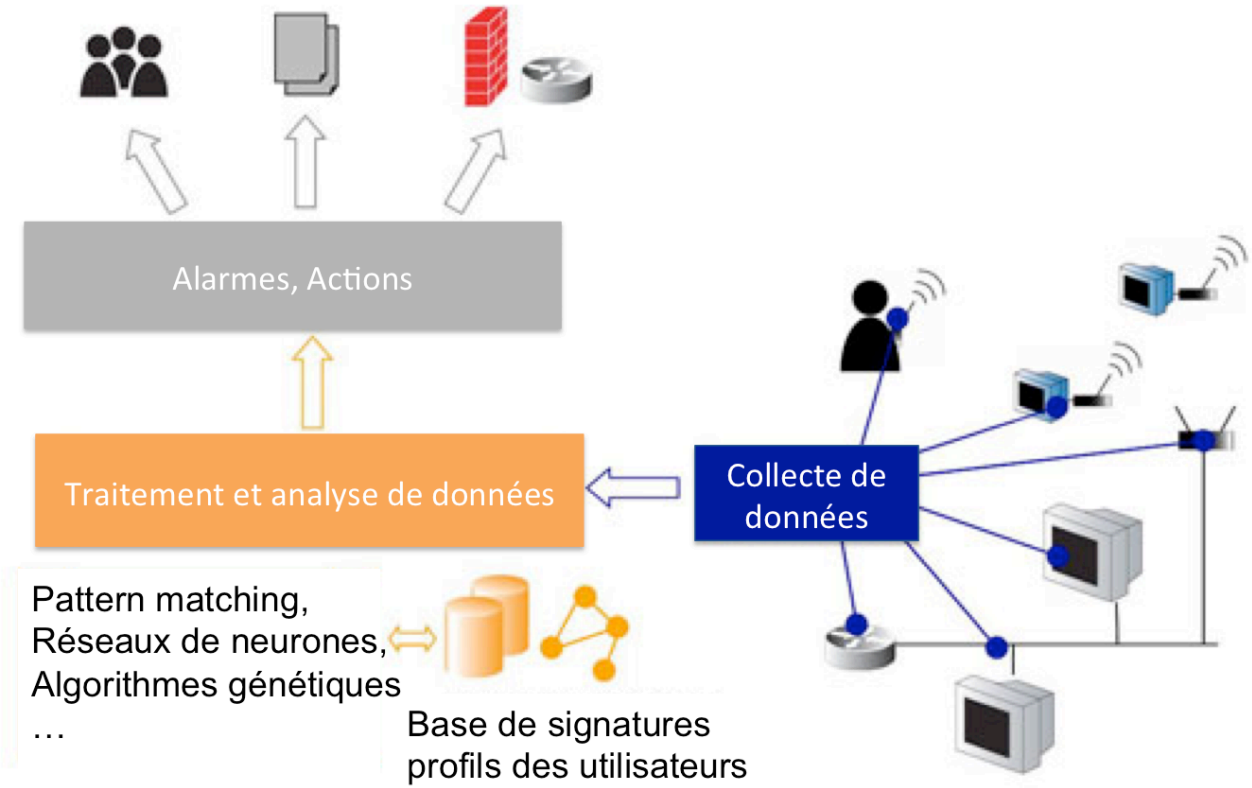
Elasticité

Possibilité d'augmenter ou de diminuer les ressources utilisées.

- La sécurité réseau devrait protéger tous les points d'accès virtuels du Cloud.
- Utiliser des procédures de sécurité pour détecter puis bloquer les menaces émergentes avant qu'elles puissent représenter un véritable danger.

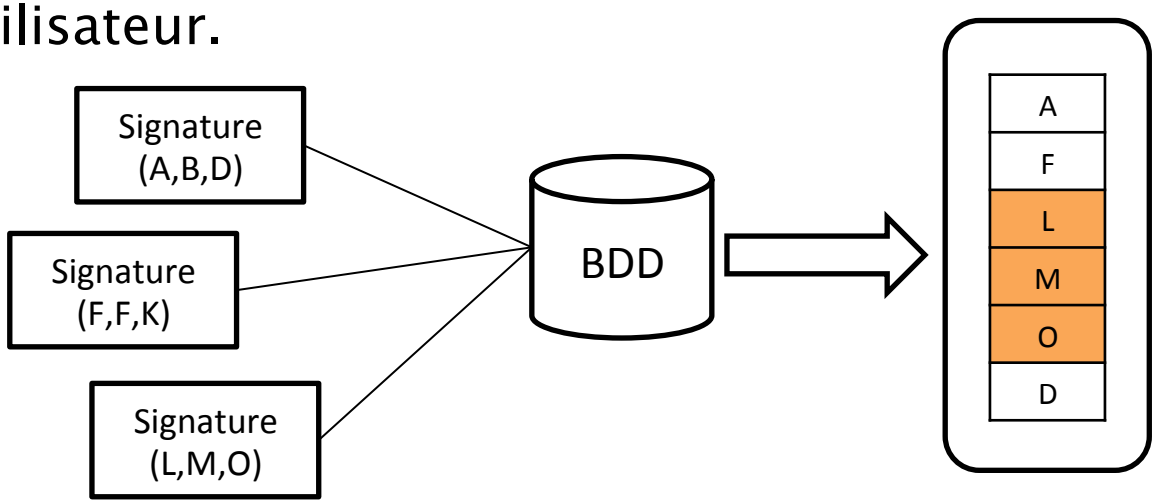
Mise en place des techniques capables de **surveiller** le Cloud pour détecter les attaques quelque soit leur origine.

IDS : un mécanisme permettant la découverte et l'analyse des activités anormales ou suspectes sur la cible analysée et de déclencher par la suite une alerte.



Approche par scénario

- Utiliser la notion des signatures d'attaques.
- Détecter les intrusions en fonction du comportement actuel de l'utilisateur.



☹ La base des règles doit être bien construite, ce qui est parfois délicat.

Systèmes experts

- Un ensemble de règles :
  - Coder ce qui est suspect (par rapport à la politique de sécurité).
  - Coder les failles et les vulnérabilités connues d'un système.
- Détecter ou pas une intrusion revient à vérifier si les informations fournies par les sondes respectent ou non les règles.

Hank S. Vaccaro and Gunar E. Liepins. Detection of anomalous computer session activity. In Proceedings of the IEEE Symposium on Security and Privacy, pages 280–289, Oakland, CA, May 1989.

Le pattern matching

Identifier dans les paquets analysés une suite d'événements caractéristiques d'une attaque :

- Le trafic réseau → chaîne de caractères principale.
- Les scénarios d'attaque → sous-suites qu'il s'agit de localiser dans cette chaîne.

Il faut disposer d'un algorithme de pattern matching efficace !

Zhang Hu. Design of Intrusion Detection System Based on a New Pattern Matching Algorithm. Computer Engineering and Technology, 2009. ICCET '09. International Conference.

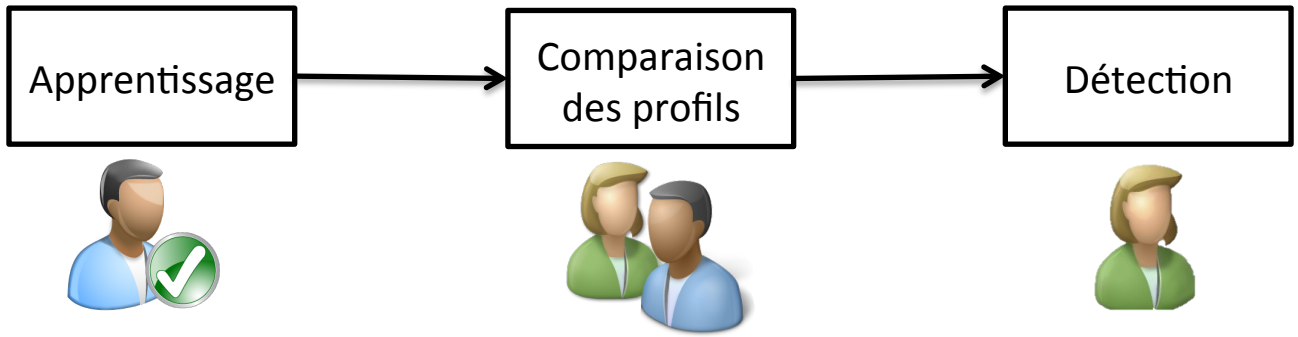
Algorithmes génétiques

- Récupérer des séries de données contenant toutes les informations nécessaires pour créer des règles.
  - IP source, IP destination, port source, port destination, protocole utilisé, etc.
- Faire correspondre les connexions à des règles déjà établies.
- Déterminer si une règle est « bonne » ou « mauvaise ».
- Les mauvaises règles sont remplacées par des nouvelles règles issues de fusions de « bonnes » règles.

Owais,S.; Snasel,V.; Kromer,P.; Abraham. Survey: Using Genetic Algorithm Approach in Intrusion Detection Systems Techniques. A Computer Information Systems and Industrial Management Applications, 2008. CISIM '0

Approche comportementale

- Détecter les intrusions en fonction du comportement passé de l'utilisateur.
- Définir un comportement normal du système (profil) et chercher ce qui ne rentre pas dans ce comportement.



☺ Détection de nouvelles attaques.

☹ Un utilisateur peut habituer le système à un comportement intrusif.

Approches statistiques

- Quantifier toute une série de paramètres liés à l'utilisateur.
  - Le taux d'occupation de la mémoire, l'utilisation des processeurs, la valeur de la charge réseau, le nombre d'accès à l'Intranet par jour, etc.
- Déterminer si la valeur générée par la sonde est normale ou non par comparaison avec le profil.
- Plusieurs modèles sont proposés :
  - Moyenne, écart-type, covariances, Markov, série temporelle, etc.

D.E. DENNING. « An Intrusion-Detection Model ». IEEE transaction on Software Engineering, 13(2): 222–232, 1987.

Réseau de neurones

Faire la différence entre ce qui est normal et ce qui ne l'est pas.

- Apprendre les séquences de commandes usuelles à chaque utilisateur.
- Prédire après chaque commande passée la commande suivante.
- Emettre une alerte en cas de déviation entre la prévision et la réalité.

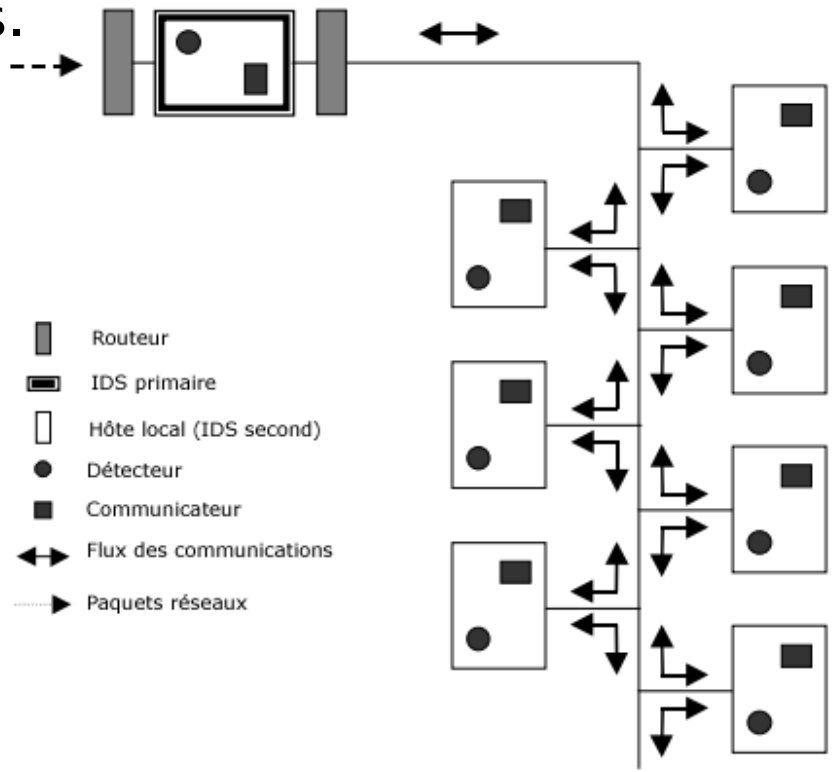
Changjun, Han Yi Lv; Dan Yang; Yu Hao. An Intrusion Detection System Based on Neural Network. Mechatronic Science, Electric Engineering and Computer (MEC), 2011 International Conference.

Réseaux immunologiques

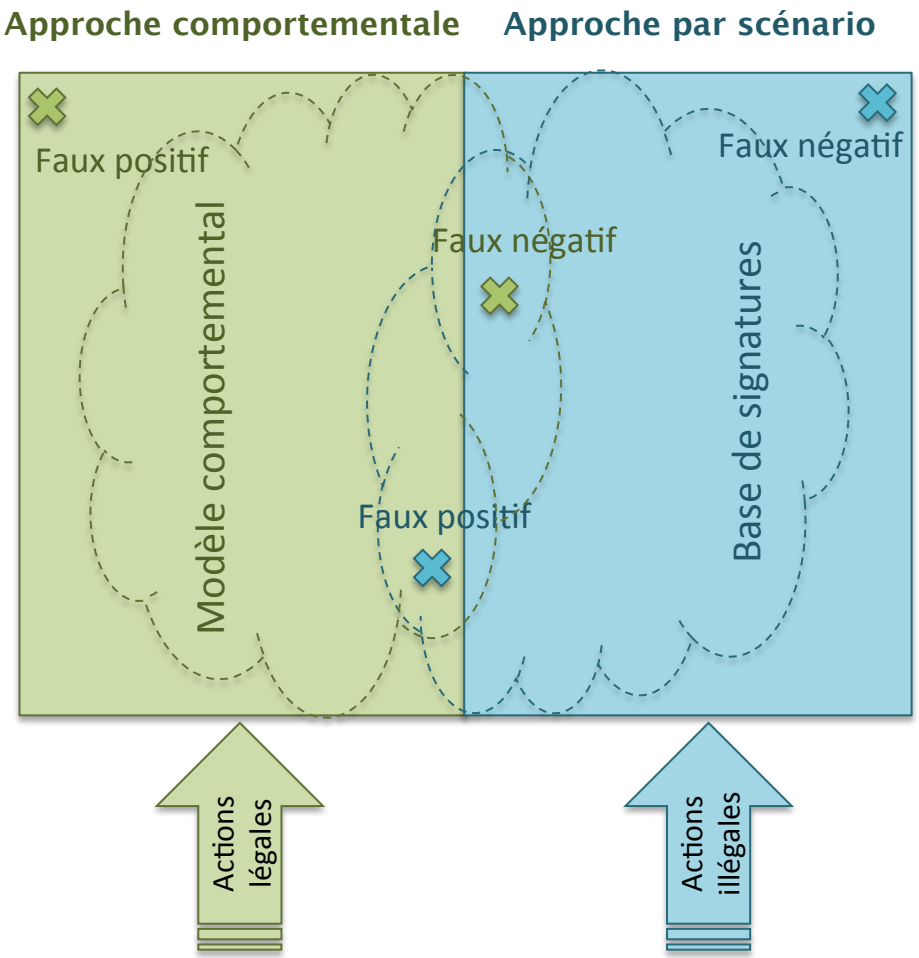
Inspirés des systèmes immunitaires.

Système immunitaire	Systèmes de détection d'intrusions
Thymus et moelle osseuse	IDS primaire (superviseur)
Nœud lymphatique	Hôte local
Anticorps	Détecteur
Antigène	Intrusion
Soi	Activité normale
Non soi	Activité anormale (suspecte)

Problématique et objectif
État de l'art
Approche proposée
Conclusion



Yan Qiao, An intrusion detection system based on immune mechanisms, SPIE Newsroom, 2007.



Approche par scénario

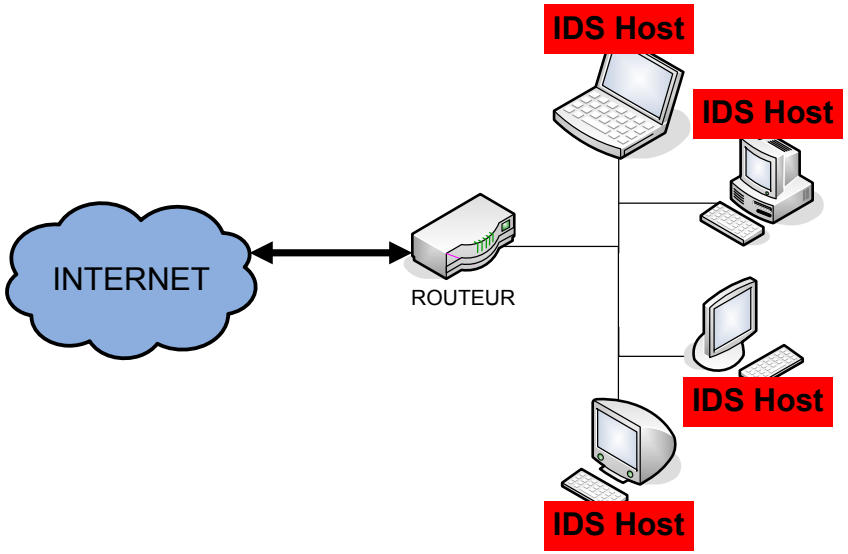
- Plus la base de signatures est mise à jour moins on aura de faux négatifs.
- Il faut connaître tous les scénarios d'attaques pour avoir moins de faux positifs.

Approche comportementale

- Plus le temps d'apprentissage est important moins on aura de faux positifs.
- Il faut s'assurer qu'un utilisateur malveillant ne puisse habituer le système à un comportement intrusif.

Couche SE : H-IDS (Host-based Intrusion Detection System)

- Analyser le trafic sur une machine.
- Capturer les paquets réseaux entrant/sortant de l'hôte.

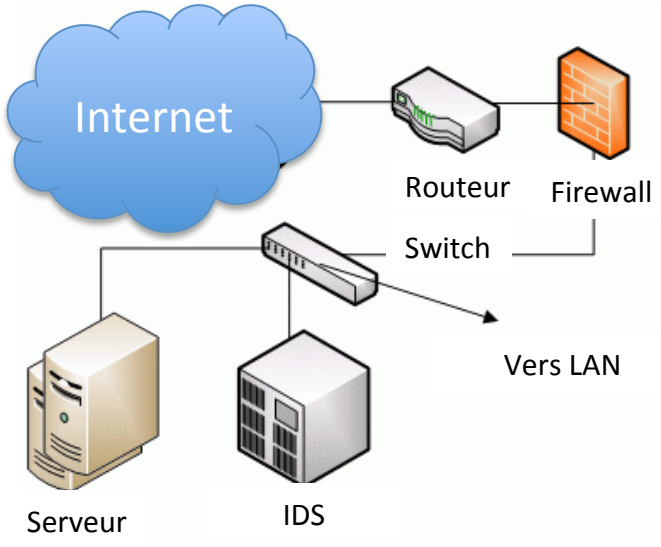


Cloud : Virtualisation

- Où mettre les HIDS ?
  - Sur chaque VM ?
  - Sur le serveur contenant les VMs ?

Couche réseau : N-IDS (Network-based Intrusion Detection System)

- La détection se fait par analyse de flux et transite sur le réseau
- Ce trafic sera ensuite analysé afin de détecter les signatures d'attaques ou les différences avec le fonctionnement de référence.



Cloud : Dématérialisation

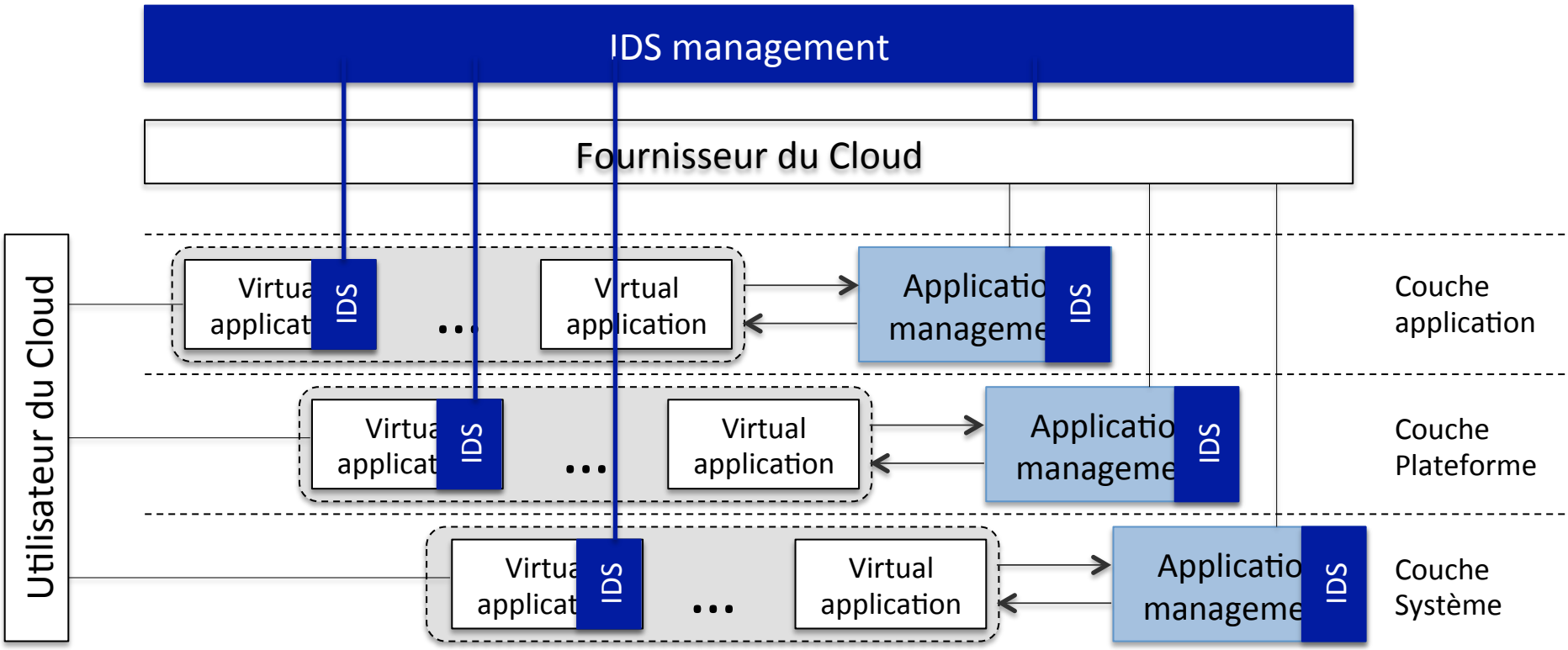
- Comment (par qui) superviser le trafic entre les VMs ?

Dans le cadre de la migration vers le Cloud Computing, la situation se complexifie encore plus :

- Tout est virtuel et immatériel.
  - Comment superviser le trafic entre deux VM ?
- Le nombre d'instances de VM tournant simultanément peut être énorme.
  - Le nombre d'évènements de sécurité générés par les IDS peut être important !
- Possibilité d'allouer rapidement des ressources supplémentaires.
  - Faut-il reconfigurer à chaque fois son IDS ?

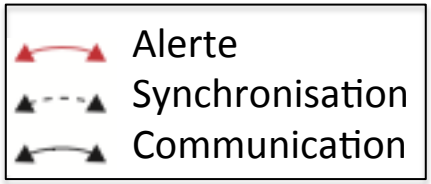
Comment déployer un IDS dans un environnement de Cloud Computing !

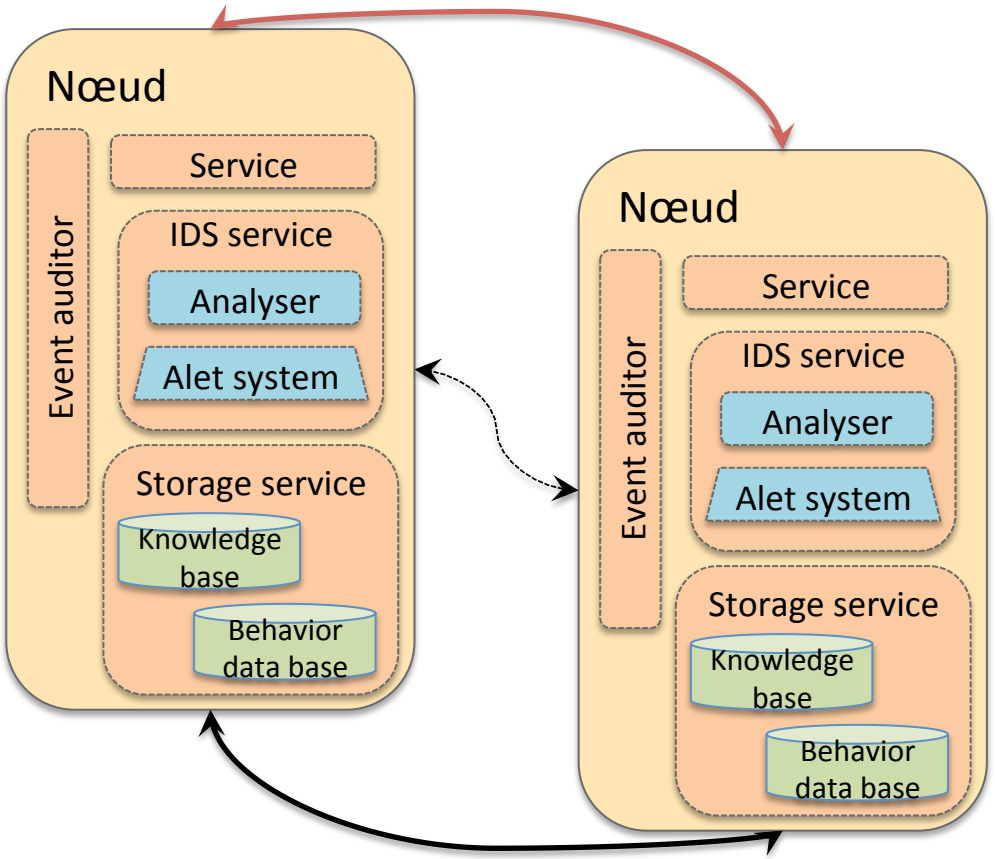
Network-based IDS



S. Roschke, C. Feng, and C. Meinel, "An Extensible and Virtualization Compatible IDS Management Architecture," Fifth International Conference on Information Assurance and Security, vol. 2, 2009

Host-based IDS





M.VIEIRA, A.SCHULTER, "Intrusion/detection techniques in grid and cloud computing environment," IEEE IT Professional Magazine, 2010.

Hypervisor-based IDS:

Un système de détection d'intrusions créé pour les hyperviseurs.

La détection se fait par :

- analyse des activités dans les machines virtuelles;
- analyse des flux et transite entre machines virtuelles et aussi entre l'hyperviseur et les machines virtuelles.

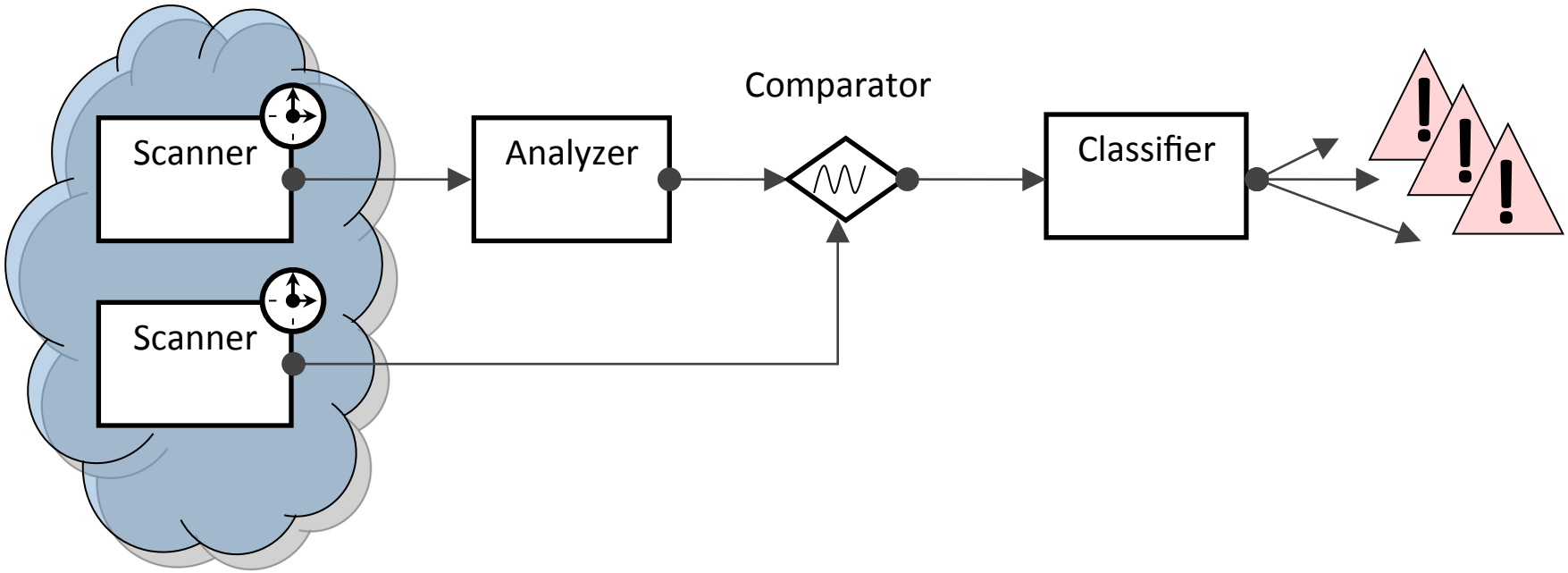
☹ Pas de documentation.

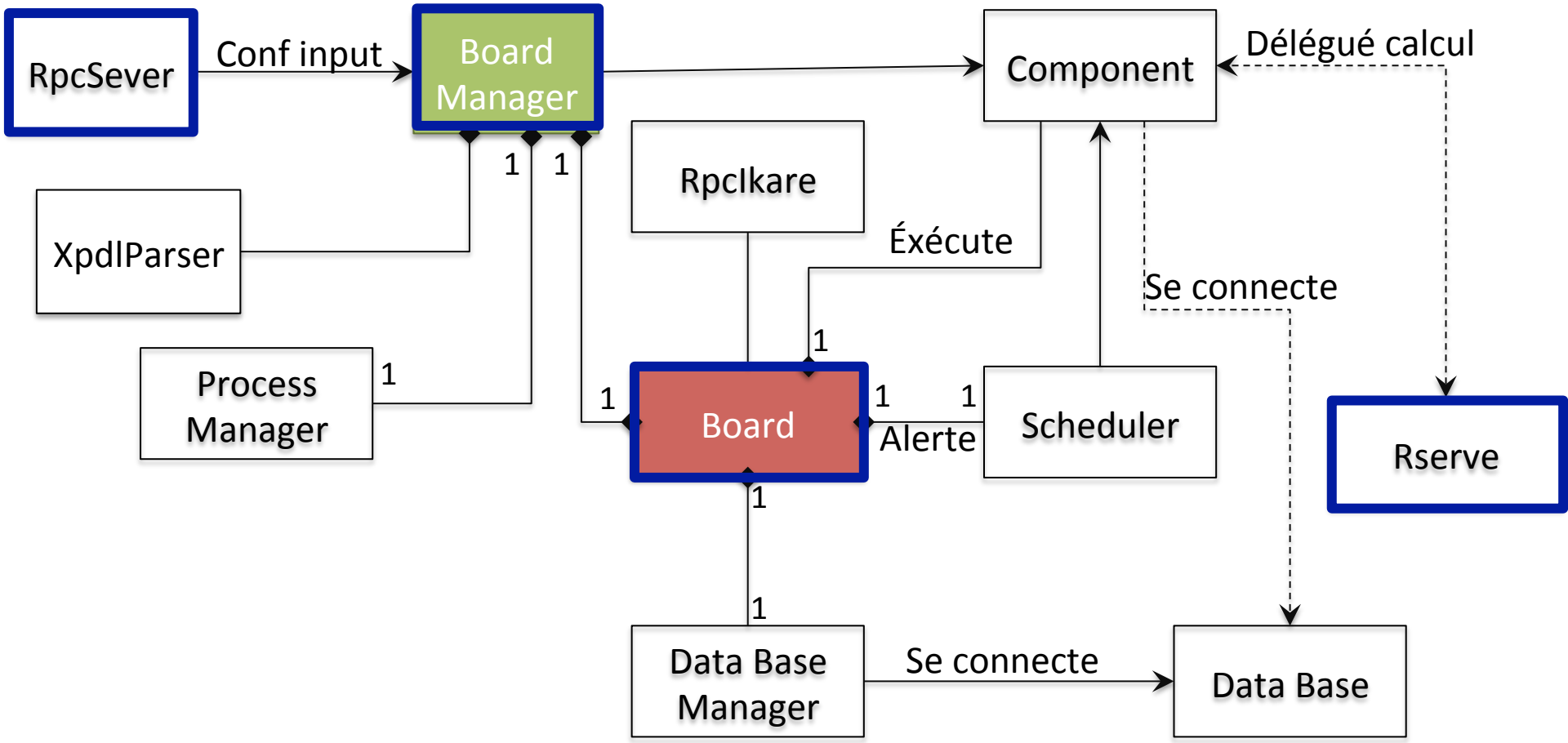
	Caractéristiques	Limites	Position dans le Cloud	Déploiement et monitoring
Host-based IDS	- Identifier les intrusion par contrôle de l'activité du système (Logs, système de fichier, etc.)	- Doit être installé sur chaque machine (VMs, Hyperviseur, hôte) - Surveille que la machine là où il est déployé	- Dans chaque machine virtuelle, dans l'hyperviseur ou dans la machine hôte	- Dans les VMs: l'utilisateur du Cloud - Dans l'hyperviseur ou la machine hôte: le fournisseur du Cloud
Network-based IDS	- Identifier les intrusion par contrôle du trafic réseau	- Difficultés de détecter les intrusions dans un réseau virtuel Ne peut détecter que des intrusions provenant du réseau où il est déployé	- Dans un réseau virtuelle ou externe	- Le fournisseur du Cloud
Hypervisor-based IDS	- Surveiller et analyser la communication entre les VMs, entre l'hyperviseur et les VMs et entre l'hyperviseur et le réseau virtuel - Le plus adapté au Cloud	- Peu de documentation - Difficile à comprendre	- Dans l'hyperviseur	- Le fournisseur du Cloud

- **iTrust (produit iKare)** : Audit en temps quasi réel, avec un moteur d'analyse des données collectées, capable de détecter les comportements déviants et d'aider à la décision.

- **VORTEX** : Adapter les compétences en optimisation et en simulation comportementale aux remontées des sondes situées dans l'infrastructure afin de déclencher des alertes de sécurité.

Mise en place du Framework PERDIX : Workflow de traitement.



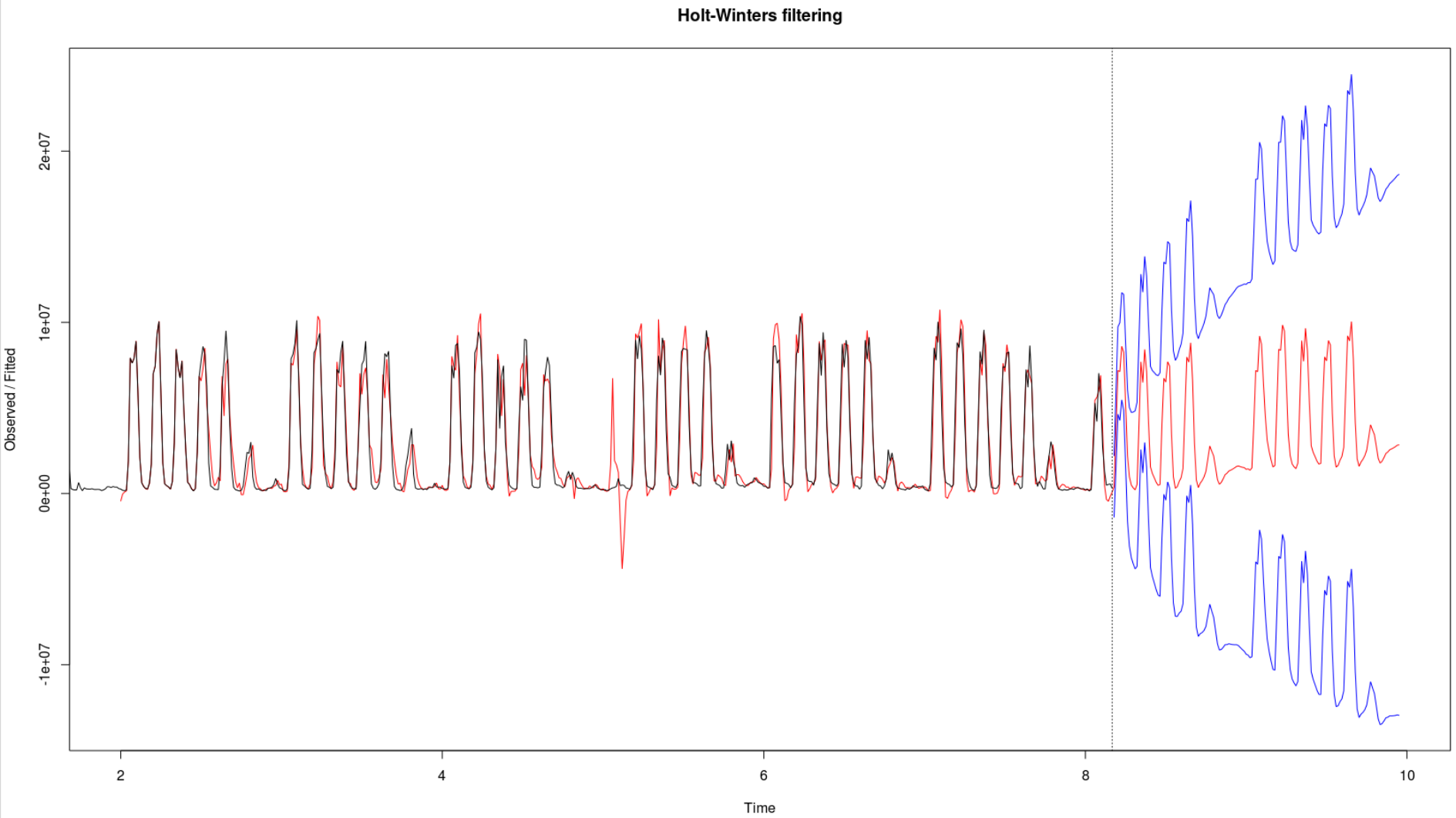

RpcSever
Conf input
Board Manager
Component
Délégué calcul
RpcIkare
XpdlParser
Éxécute
Se connecte
Process Manager
Board
Scheduler
Alerte
Rserve
Data Base Manager
Se connecte
Data Base
1

Scanner

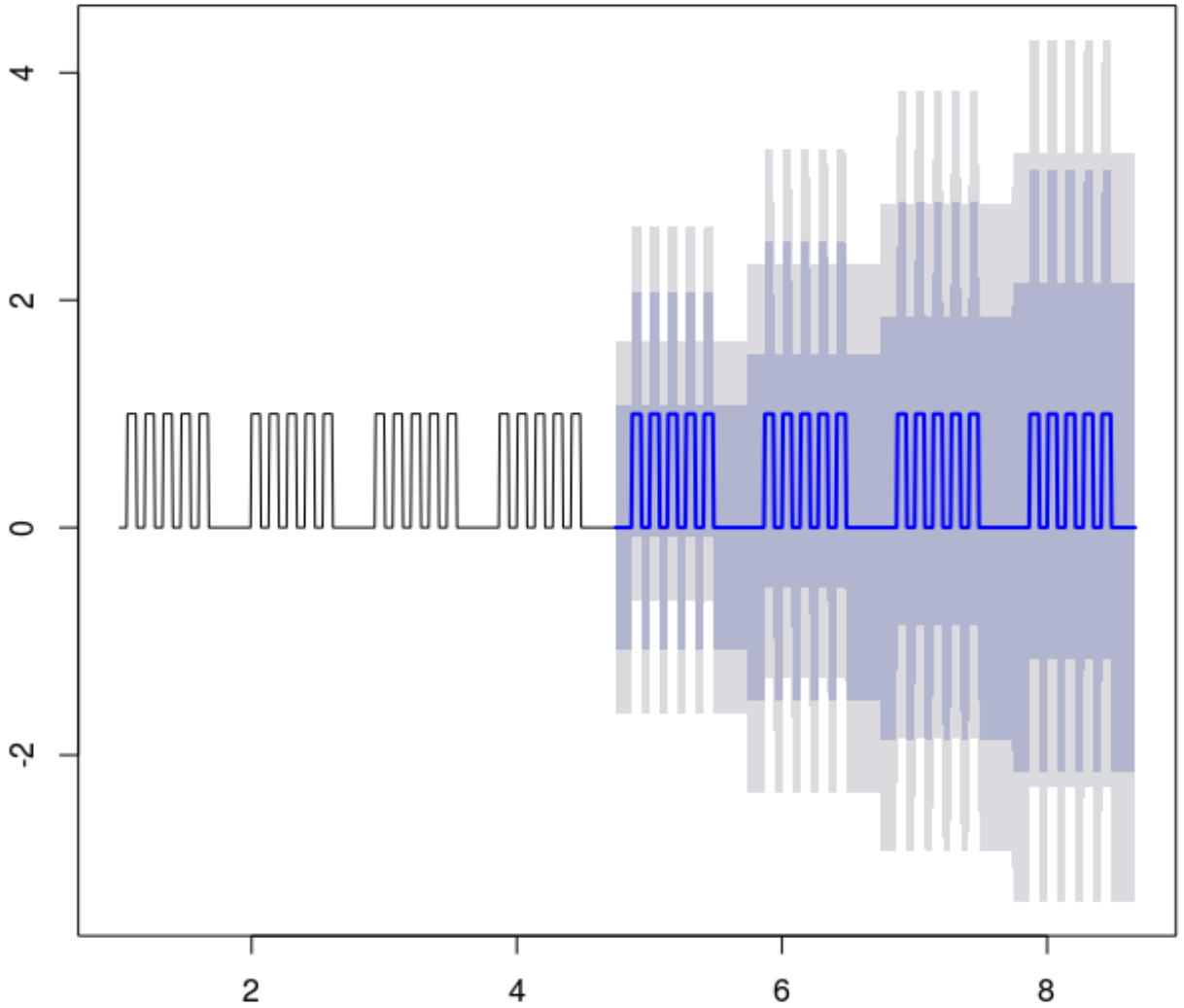
- QoS :
  - TCP-PING : évaluer le temps de connexion sur les ports ouverts;
  - HTTP/HTTPS : évaluer le temps de réponse d'un serveur selon les 4 temps (dns, connect, send, receive).
  - FTP/SFTP.
- SNMP (optionnel suivant le host ou l'équipement réseau) : obtenir les valeurs instantanées du trafic ethernet IN/OUT, la consommation CPU et l'utilisation de la mémoire.

Analyser

- Les informations sont réduites par l'enchainement de divers algorithmes mathématiques :
  - Utilisation de la transformée de Fourier rapide (FFT).
  - Test de Kolmogorov Smirnov.
  - Utilisation de la méthode Holt-Winter pour la recherche d'une saisonnalité dans le flot des données continues;
  - Utilisation des approches statistiques : moyenne, variance et écart-type.
- Prédiction d'états :
  - Modèle ARIMA.
- Nécessité de corrélation.

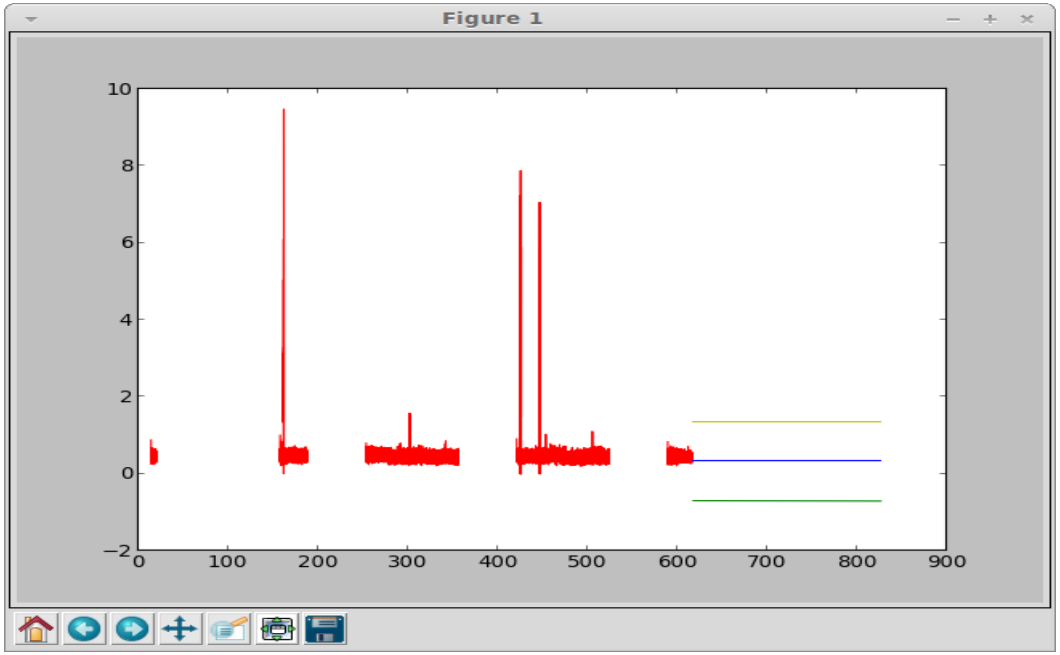

Holt-Winters filtering
Observed / Fitted
2e+07
1e+07
0e+00
-1e+07
2
4
6
8
10
Time

Prédiction d'états (ARIMA)



Générer des alertes

- Générer des alertes quand la valeur prédite est différente de la valeur mesurée (dans un intervalle de confiance).
- La sensibilité peut être réglée à travers des seuils configurables.



Conclusion
Reste à faire

- Scanner / récupération des données.
- Analyse des données.
  - ☑ Niveau réseau.
  - ☑ Niveau système d'exploitation.
  - ☑ Réduction de donnée.
  - ☑ Approche comportementale.
    - ☑ Méthodes statistiques : Holt-winter, moyenne, écart-type, variance.
  - ☑ Corrélation.
- Levée d'alertes.
- Expérimentations.

- Tester avec un nombre plus grand de machines.
- Améliorer (Ajouter) les algorithmes.
- Utiliser un mécanisme d'apprentissage et de classification pour produire un état de sécurité du système.
- Gérer les alertes et minimiser les fausses alarmes.

MERCI
pour votre attention