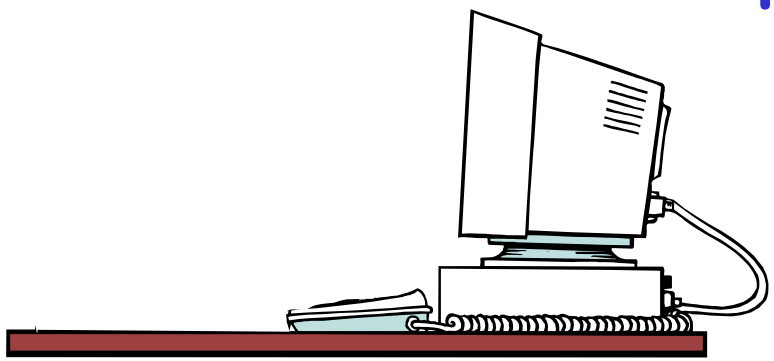
Détection d'intrusions : mythes passés, réalité présente, défis futurs



Marc Dacier
Institut Eurecom

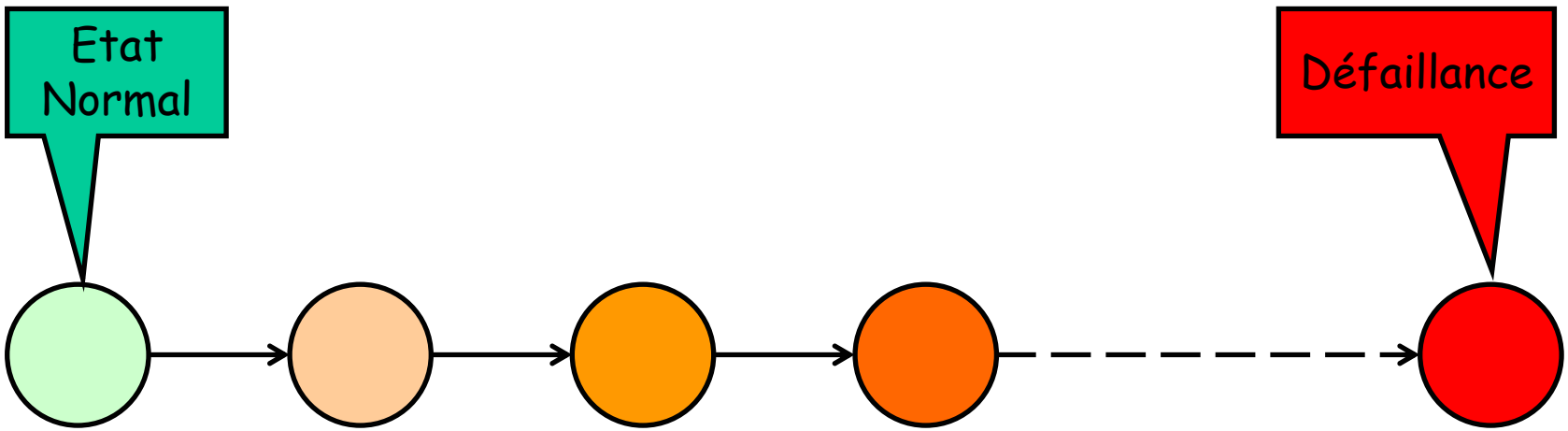
http://www.eurecom.fr/~dacier
marc.dacier@eurecom.fr

Plan général

- Introduction :
  détection d'intrusions et sûreté de fonctionnement
- Mythes Passés :
  les travaux fondateurs du domaine
- Réalité Présente :
  taxonomie
  le monde « réel »
  la recherche
- Défis futurs :
  Formalisation et validation

Attributs de la SdF

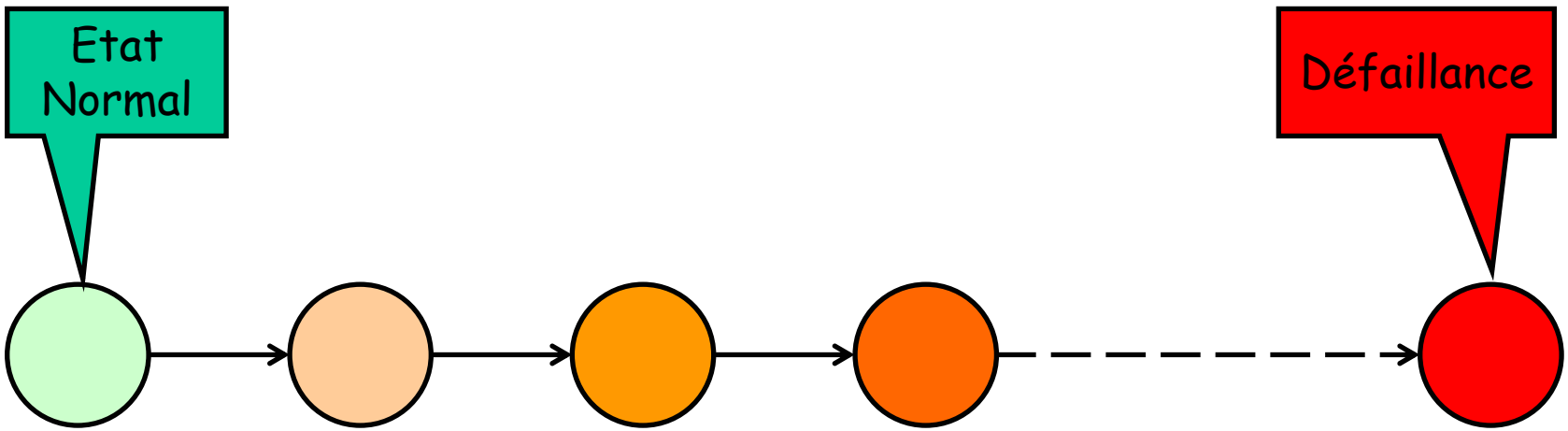
- Disponibilité
- Fiabilité
- Sécurité-innocuité
- Confidentialité
- Intégrité
- Maintenabilité



Guide de la sûreté de Fonctionnement, ouvrage collectif du LIS, sous la direction de J-C Laprie, CEPADUES Editions, 2.85428.341.4

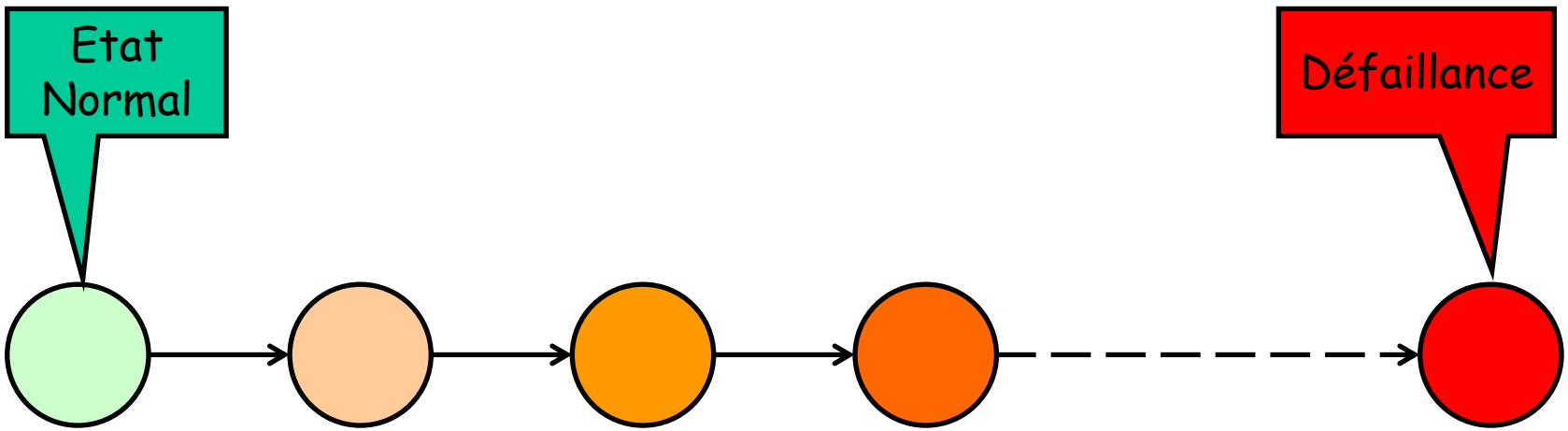
Moyens de la SdF: Prévention

- Prévention de fautes: « comment empêcher l'occurrence ou l'introduction de fautes »
  - Bonnes pratiques d'ingénierie
  - Pare feu, authentification forte, contrôle d'accès, etc …



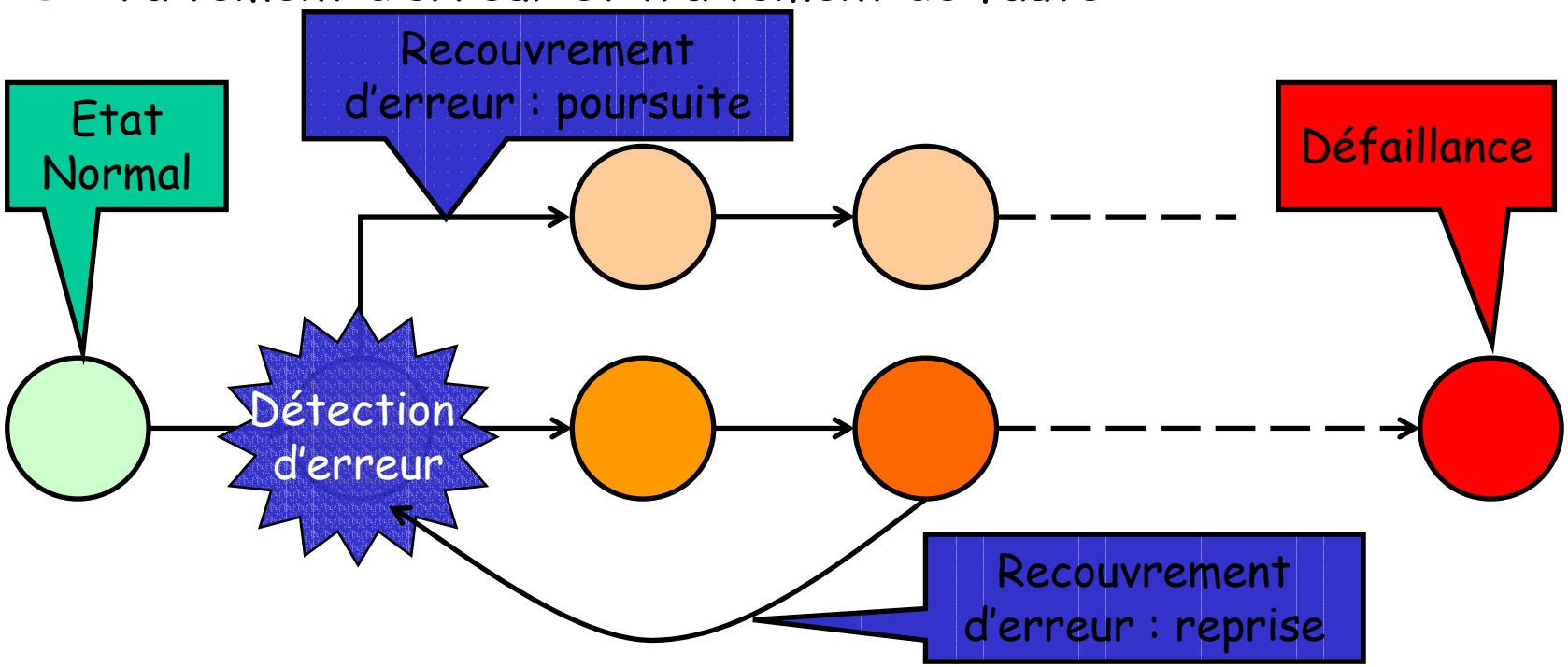
Moyens de la SdF: Elimination

- Elimination de fautes: « comment réduire (nombre, sévérité) des fautes »
  - Vérification statique, vérification dynamique (test)
  - Tests de pénétration, vérification de configurations, etc …



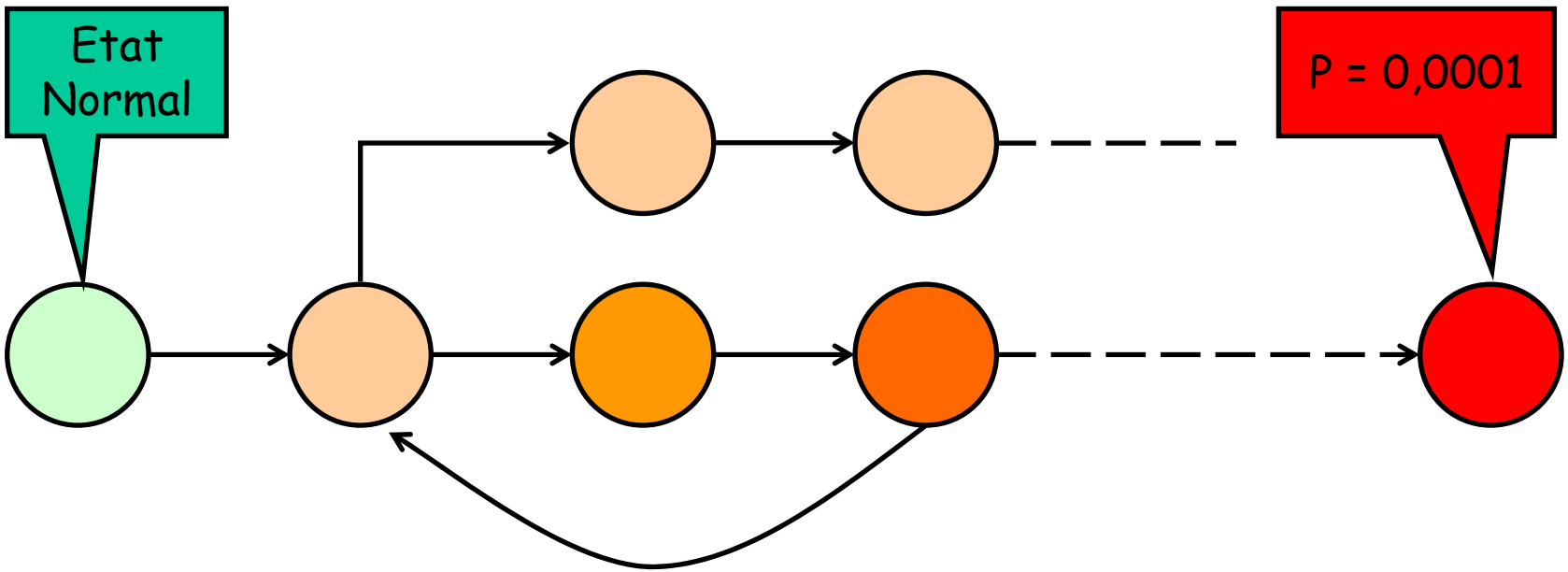
Moyens de la SdF: Tolérance

- Tolérance aux fautes: « comment fournir un service à même de remplir la fonction du système en dépit des fautes »
  - Traitement d'erreur et traitement de faute



Moyens de la SdF: Prévision

- Prévision de fautes: « comment estimer la présence, la création et les conséquences des fautes »
  - Evaluation ordinale, evaluation probabiliste
  - « analyse quantitative des risques »



Tolérance aux fautes

- Traitement d'erreur :
  - **Détection d'intrusions**
  - Diagnostic d'erreur
  - Recouvrement d'erreur
- Traitement de faute:
  - Diagnostic de faute
  - Passivation des fautes

Erreurs dues à des fautes humaines, malicieuses, externes, opérationelles, temporaires, etc ….

Origine des vulnérabilités

- Les utilisateurs !!!
- La très grande difficulté d'identifier toutes les vulnérabilités
- L'impossibilité de retirer toutes les vulnérabilités connues
- Le choix raisonné et conscient de ne pas retirer toutes les vulnérabilités connues

Origine des attaques

- Espionnage Industriel
- Robots
- Cyber terroristes
- Script Kiddies

Plan général

- Introduction :
  détection d'intrusions et sûreté de fonctionnement
- Mythes Passés :
  les travaux fondateurs du domaine
- Réalité Présente :
  taxonomie
  le monde « réel »
  la recherche
- Défis futurs :
  Formalisation et validation

Les mythes du passé

Y'A QU'A

FAUT QU'ON

T' AS QU' A

Les mythes du passé

- En 1980, James P. Anderson propose :
  « changes to computer audit mechanisms to provide information for use by computer security personnel when tracking problems »
- Il introduit la notion de « audit reduction »
- .. Et l'utilisation de
  « some sort of statistical analysis of user behavior [...that...] might represent a way of detecting masqueraders »

IDES/NIDES

- 1984-1986: D. Denning et P. Neumann
- IDES: Intrusion Detection Expert System
  - US Navy's Space and Naval Warfare Systems Commands (SPAWARS)
- Approche comportementale
- Mise à jour continue des profils
- Un peu d'approche basée sur la connaissance
- NIDES = IDES en réseau

A Real Time Intrusion Detection Expert System (IDES), SRI Project 6784, Final Technical Report, Feb. 28, 1992

System Design Document: Next Generation Intrusion Detection Expert System (NIDES), March 9, 1993

IDES: les mesures

- Intensité d'activité:
  - Nombre d'événements dans un intervalle de temps (pour détecter les pics)
- Distributions des records d'audit
  - Comparaison du passé proche à l'historique
- Mesures par catégorie:
  - Comparaison de la distribution des noms de fichiers, utilisateurs, etc .. p. r. à l'historique
- Mesures absolues
  - Comparaison de la distribution des valeurs de compteurs (cpu time, I/O, etc.) p. r. à l'historique

Audit Analysis Project

- 1984-1985
- L. Halme, T. Lunt, J. van Horne
- Sytek (pour US Navy's SPAWARS command)
- Collecte de données au niveau du shell Unix
- Analyse off line par le biais de requêtes sur une SGBD
- Offre une première validation de l'approche comportementale

« Automated Analysis of Computer System Audit Trails for Security Purposes », L; Halme, T. Lunt, J. Van Horne, Proc. Of Nat. Comp. Security Conf., Washington, D.C Sept. 1986

Discovery

- 1986: environnement commercial de collectes de données
- Détection d'intrusions au niveau applicatif
- But:
  - Détection des accès non autorisés
  - Utilisation malicieuse de privilèges
  - Transactions invalides dans une base de données de crédit en ligne
- Détection off line sur base de méthodes statistiques
- (Cobol, IBM 3090)

Discovery, an Expert System in the Commercial Security Environment », W. Tenner, Proc. Of IFIP Security Conf., Monte Carlo, 1986

MIDAS

- Multics Intrusion Detection and Alerting System (1989 – mid 1990's)
- NCSC, système de haute sécurité
- Analyse des données d'audit, enrichies avec des données venant d'autres sources
- Système hybride: règles + profils
- Effet de bord: intimidation
- (LISP, station de travail Symbolics)

« Expert Systems in Intrusion Detection: A Case Study », M. Sebring et al., Proc. Of the Eleventh Nat. Comp. Sec. Conf., Washington D. C., Oct. 1988

NADIR

- Network Audit and Intrusion Reporter
- Los Alamos Nat. Lab., K. Jackson
- Off line, basé sur un SGBD Sybase
- Système hybride: règles + profils
- La topologie réseau est prise en compte
- Utilisé durant de nombreuses années
- Très efficace pour détecter de nouveaux utilisateurs !
- Très coûteux en ressources et en temps

« NADIR, An Automated System for Detecting Network Intrusion and Misuse: A Case Study », Comp. And Sec. 12, N.3, May 1993, pp. 235-2348

Network System Monitor (NSM)

- Première approche basée réseau
- Univ. Of California Davis, 1989
- Les paquets sont analysés par rapport à:
  - Des profils statistiques
  - Des règles pré établies de haut niveau
- Un test sur site réel durant deux mois révéla 300 attaques.
- Les administrateurs n'en avaient remarqué qu'un pourcent ….

« Detection of Anomalous Computer Session Activity », H. S. Vaccaro, G. E. Liepins, Proc. Of the 1989 IEEE Symp. On Res. In Sec. And Privacy, Oakland, CA, May 1989, pp. 280-289

Et l'Europe ??

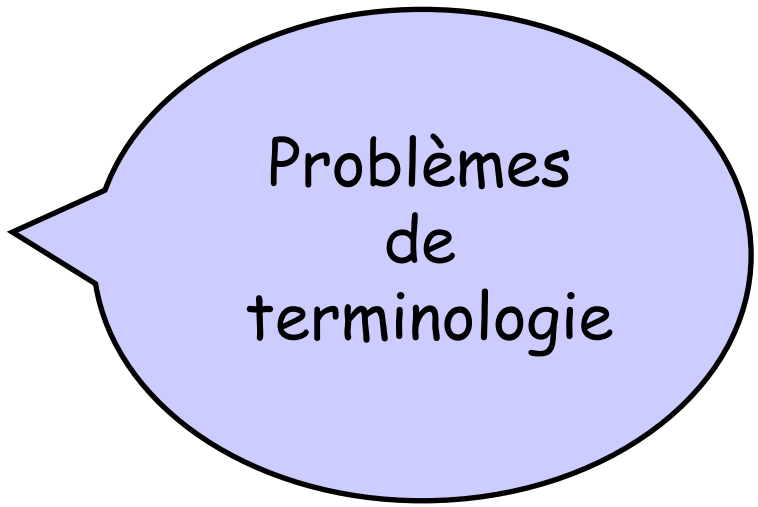
- Securenet: projet européen visant à construire la première architecture intégrant plusieurs approches.
- Premiers travaux en 1992 sur l'utilisation des réseaux neuronaux (Debar et al.)
- Premiers travaux sur l'utilisation des algorithmes génétiques (Mé et al.)
- Premiers travaux sur la détection dans les domaines des télécommunications GSM (Molva et al.)
- Premier langage spécifique d'expression des règles de détection (ASAX).
- ....

Plan général

- Introduction :
  détection d'intrusions et sûreté de fonctionnement
- Mythes Passés :
  les travaux fondateurs du domaine
- Réalité Présente :
  taxonomie
  le monde « réel »
  la recherche
- Défis futurs :
  Formalisation et validation

Détection d'intrusions aujourd'hui

- Traitement d'erreur :
  - Détection d'erreur
  - Diagnostic d'erreur
  - Recouvrement d'erreur
- Traitement de faute:
  - Diagnostic de faute
  - Passivation des fautes



Taxonomie (vue partielle)

- Paradigmes
- Utilisation
- Comportement
- Sources de données

Paradigmes

- Approche comportementale

- Approche basée sur la connaissance

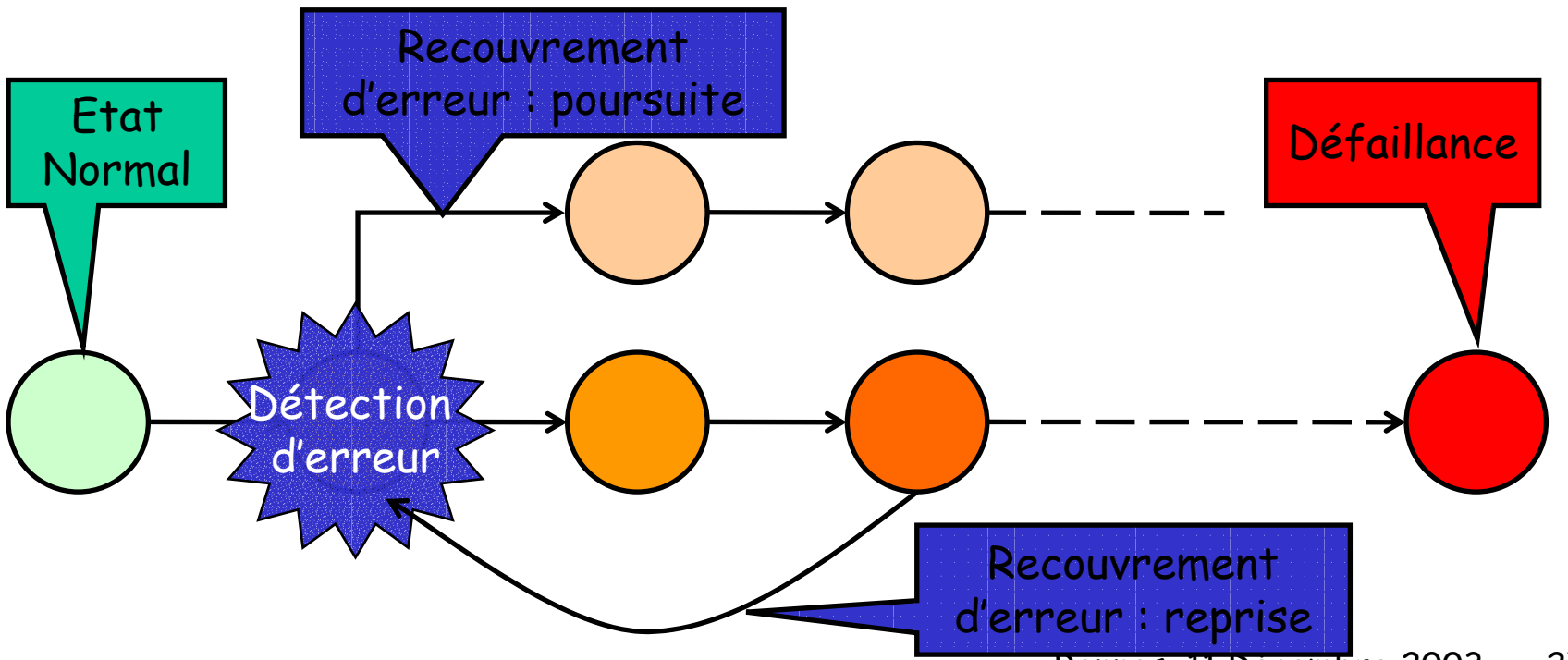
Utilisation

- Continue
  - Permet une réaction en temps réel
  - Nécessite des ressources dédiées

- Sporadique
  - Ne permet que de voir les traces laissées par des attaques passées
  - Plus simple à mettre en oeuvre

Comportement

- Passif : recouvrement d'erreur par poursuite
- Actif: recouvrement d'erreur par reprise



Sources de données

- Données disponibles sur l'hôte cible:
  - Fichiers d'audit, Fichiers de log applicatifs, Appels systèmes, etc.
  - la notion d'hôte cible est à prendre au sens large (ex. algorithme de routage, serveurs de noms)
- Données en transit sur le réseau :
  - Paquets TCP, IP, ICMP, etc …
  - Consommation de bande passante,
  - Flux observés
- Données contextuelles

Taxonomie (vue partielle)

- Paradigmes
- Utilisation
- Comportement
- Sources de données

Plan général

- Introduction :
  détection d'intrusions et sûreté de fonctionnement
- Mythes Passés :
  les travaux fondateurs du domaine
- Réalité Présente :
  taxonomie
  le monde « réel »
  la recherche
- Défis futurs :
  Formalisation et validation

Le monde réel

- Depuis Février 2001 (attaques DDoS), la détection d'intrusions est sortie du monde de la recherche pure
  - NB: le premier produit commercial était Français, apparu au milieu des années 90'
  - TRM, premier produit offrant une console de corrélation pour des alertes issues d'outil de détection d'intrusions (Debar et al.), sort du laboratoire IBM de Zurich en juin 2001.
- Evolutions récentes:
  1. Approches basées hôtes,
  2. Approches basées réseaux,
  3. Approches hybrides,
  4. Approches hybrides et apport contextuel
  5. « IPS »

Les problèmes rencontrés

- La complexité algorithmique inhérente aux approches basées sur la connaissance nécessite une simplification des règles de recherche des signatures d'attaques

- Conséquences : fausses alarmes positives

Les problèmes rencontrés (suite)

- Une spécialisation des règles, afin d'éliminer le premier problème, ouvre une large brèche aux méthodes d'évitement des techniques de détection

- Conséquence: fausses alarmes négatives

Solutions apportées

- Normalisation du trafic (V. Paxson)
- Postprocessing des alertes, selon une approche comportementale, permet de supprimer près de 99% des fausses alertes positives (Julisch et al.)
- Optimisation des algorithmes de pattern matching
- Solutions matérielles pour traiter les hauts débits

Les solutions existantes

- Chaque jour un peu meilleures mais globalement très décevantes.
- Suite aux revers essuyés par leurs clients lors de déploiements, Gartner Group recommande de postposer tout investissement en IDs pour s'orienter plutôt vers des IPs (!?) qui ne sont rien d'autre que des IDS au comportement actif
  - Ces IPs ont besoin des IDs !!!

Tendances fortes

- Diversité de détecteurs
  - L'open source reste très populaire
- Consoles de corrélation
  - Très utiles mais le prix d'entrée est élevé
  - Évolution très forte à prévoir dans les 3 ans à venir
  - Intégrées de plus en plus aux environnements de gestion de réseaux et systèmes pour pouvoir générer des contre mesures et obtenir des informations contextuelles
- « Out tasking » des services de détection d'intrusions

Plan général

- Introduction :
  détection d'intrusions et sûreté de fonctionnement
- Mythes Passés :
  les travaux fondateurs du domaine
- Réalité Présente :
  taxonomie
  le monde « réel »
  la recherche
- Défis futurs :
  Formalisation et validation

Les « indémodables »

- Langages et moteurs d'inférence:
  - Pour les signatures d'attaques
  - Pour les règles de corrélation

- DoS et DDoS, Port Scan, vers :
  - Du fait de l'absence de données réelles, les modèles proposés sont parfois en décalage complet avec la « réalité » telle que supposée par les experts.
  - Pour les mêmes raisons, le coût des solutions préventives architecturales ne peut être estimé de façon réaliste.

Une branche très active

- Modélisation du comportement de processus:
  - Travail initial par Forrest et al.,
  - Repris et amélioré par Wespi et al.
  - Puis par de nombreux autres groupes pour évaluer et améliorer la complexité algorithmique et les risques d'évitement (DFA, NDFA, PDA, DPDA, Dyck Model, …)

Des approches 'proactives '

- Idée: trouver les attaques avant les attaquants
  - Protocoles de routages
  - Nouvelles techniques de propagation de vers
  - Attaques de réseaux P2P
  - Techniques d'évitement de solutions connues

Une tendance plus récente

- Aux USA, le financement de la recherche dans ce domaine précis a changé.
- D'origine principalement militaire au départ, puis soutenu par la DARPA, il attirait les laboratoires privés (ex. SRI).
- La NSF tend à remplacer la DARPA dans ce domaine. Le type de recherche s'en trouve influencé.

Plan général

- Introduction :
  détection d'intrusions et sûreté de fonctionnement
- Mythes Passés :
  les travaux fondateurs du domaine
- Réalité Présente :
  taxonomie
  le monde « réel »
  la recherche
- Défis futurs :
  Formalisation et validation

Défis futurs

- Architecture de détection
- Corrélation d'alertes
- Tolérance aux intrusions
- Validation
- Aspects légaux

Architecture de détection

- Quels détecteurs choisir ?
- Où placer ces détecteurs pour :
  - Maximiser la capacité de détection
  - Minimiser les coûts de déploiement.
- Comment intégrer cette application distribuée au sein d'architectures de gestion existantes ?
- Comment choisir la 'meilleure' architecture ?

Corrélation d'alertes

- Que veut dire 'corrélation' ?
- Comment rendre la vie plus simple aux opérateurs ?
- Comment sélectionner les données les plus pertinentes ?
- Comment obtenir toute l'information utile et seulement celle là ?
- Comment tirer parti des connaissances issues d'autres domaines ?
- Comment évaluer les solutions retenues (attention aux usines à gaz) ?

Tolérance aux intrusions

- Problème: le détecteur est la première cible de l'attaquant.
- Comment construire un système tel qu'il puisse détecter qu'il est lui-même partiellement sous contrôle externe ?
- Comment évaluer (prouver ?) l'efficacité de la solution ?

Validation

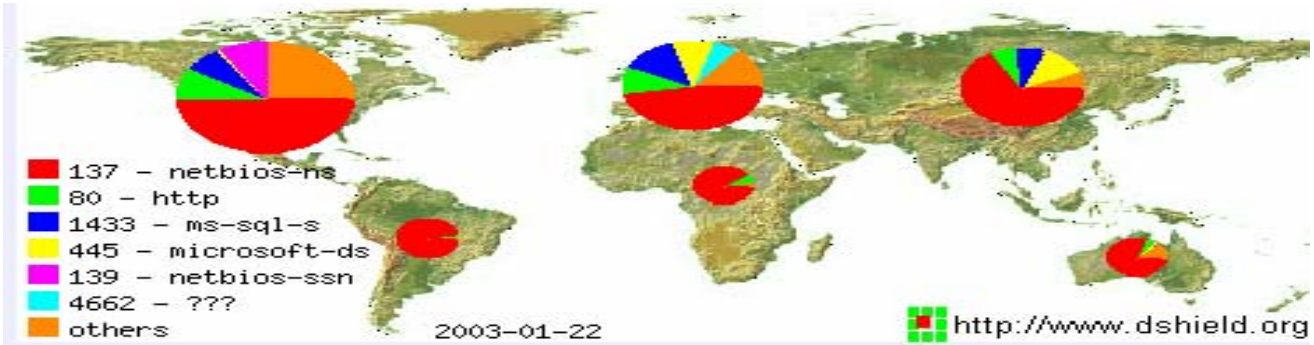
- Que sommes nous censés détecter ?
- Quelles sont nos hypothèses de faute ?
- Ces hypothèses sont elles correctes ?
- Ces hypothèses doivent elles revues dans le temps ?
- Peut on faire les mêmes hypothèses pour tous les systèmes à défendre ?

Validation (suite)

- Absence totale de données représentatives non biaisées et publiques pour valider les travaux effectués !
- Les approches basées sur des données « synthétiques » ont montré leurs défauts.
- L'anecdotique et le sensationnalisme brouillent l'horizon.
- Les données existantes sont jalousement gardées par leurs propriétaires.

Exemple

- 24 janvier 2003 (www.incidents.org)
  - Pommes et bananes



Top Attacked Ports

netbios-ns	137
http	80
ms-sql-s	1433
microsoft-ds	445
netbios-ssn	139
???	4662
domain	53
telnet	23
???	135
???	8714

Trends

2647	syncserver
37	time
4672	rfa
3456	Terrortrojan
4665	eDonkey2000
4662	eDonkey2000
4661	eDonkey2000
135	epmap
137	netbios-ns
80	www
6346	BearShare

Defacement

(http://www.attrition.org/mirror/attrition/2001/07/13/www.sans.org/)

Fluffi Bunni ownz you.

A BamBam here a dotslash there
here a dot there a slash
everywhere a dot slash

look mommy im on sans !

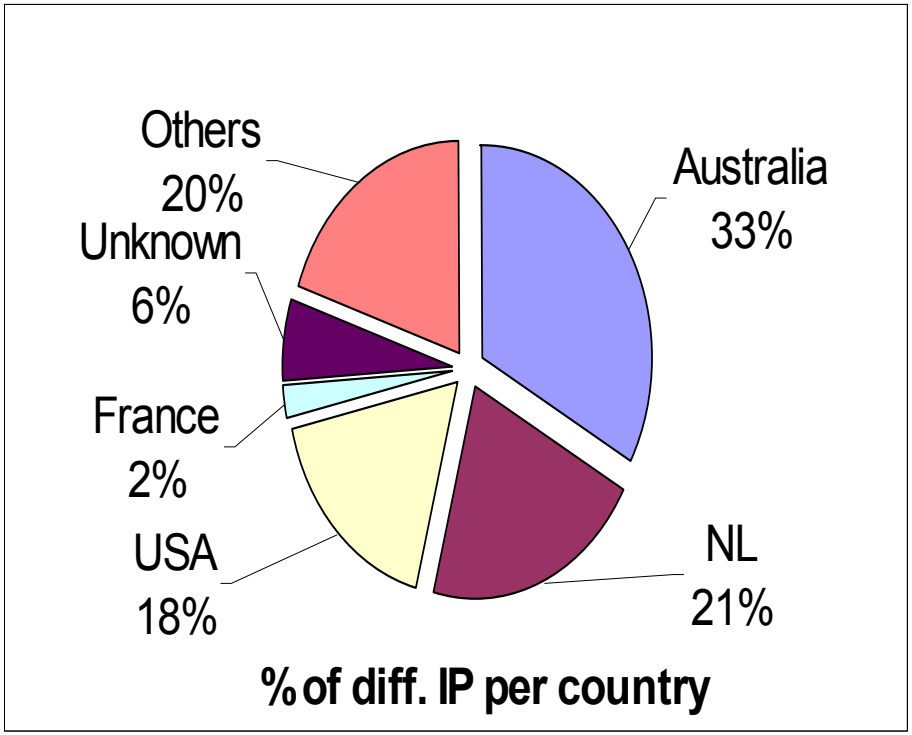
Une piste intéressante

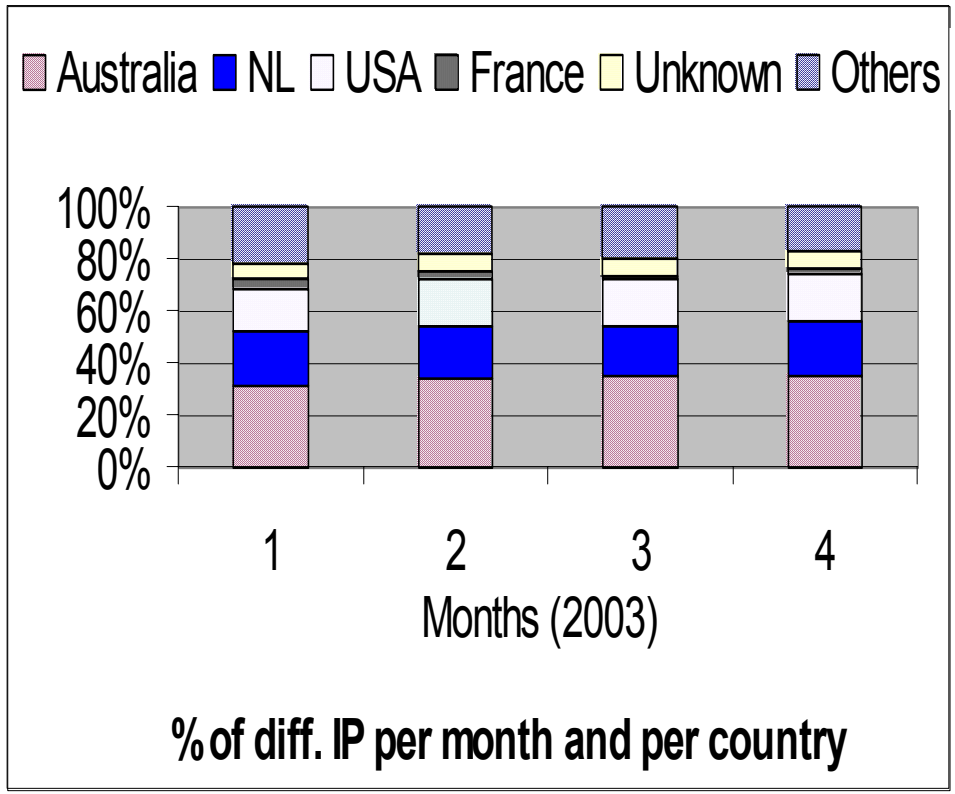
- Utilisation de « pots de miel »
  - Machines « sacrificielles » destinées à collecter des informations
- Des logiciels existent mais des plateformes collaboratives de collecte distribuées n'existent toujours pas !
- Pour découvrir le domaine: www.honeynet.org

Quelques résultats

Collaboration avec FT R&D (H. Debar)

- Nous ne sommes pas tous soumis aux mêmes menaces



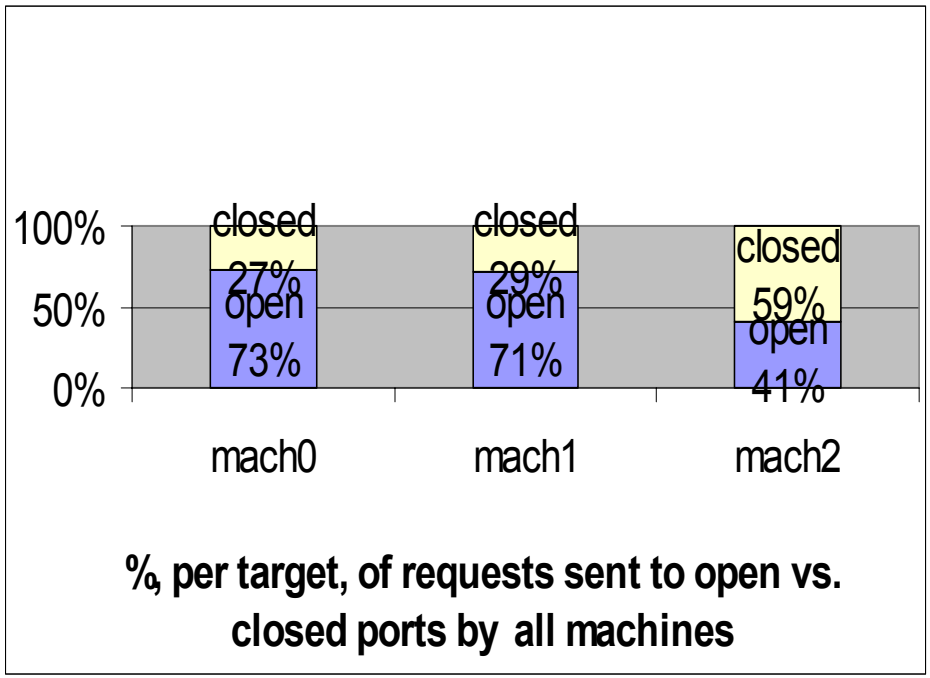


Quelques résultats

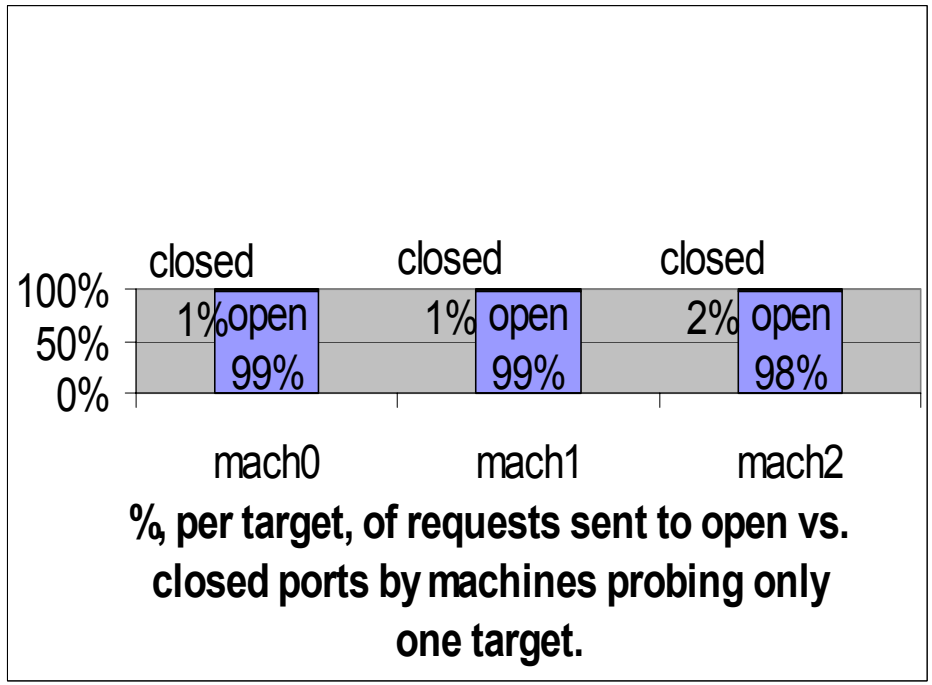
Collaboration avec FT R&D (H. Debar)

- Le processus d'attaque est un processus collaboratif

  L'analyse complète sera publiée à PRDC-10 (février 2004)



%, per target, of requests sent to open vs. closed ports by all machines



%, per target, of requests sent to open vs. closed ports by machines probing only one target.

Invitation

- L'institut Eurecom est en train de mettre sur pied une large plate forme distribuée de collecte d'informations basée sur des pots de miel.

- Nous sommes intéressés par:
  - Des partenaires désireux de déployer un honeypot chez eux selon un mode opératoire très précis et dont les données seront visibles par tous les participants à la plateforme.
- N'hésitez pas à me contacter …

Invitation (suite)

- RAID 2004 aura lieu à Sophia Antipolis (Nice) du 15 au 17 septembre 2004.
- Organisé par l'institut Eurecom (R. Molva general chair).
- Co localisée avec ESORICS 2004 du 13 au 15
- Deadline pour le CFP: 31 Mars 2004

Pour plus d'informations :

dacier@eurecom.fr