Définitions & Menaces

- Principaux risques
- Quelques types d’attaques
- Système d’Information (SI)
- 6 principes de défense
- Théorie de l’information (rappel)

Principaux risques

- Personnel

  Perte de données, vol de données (numéro de carte bancaire), …

- Entreprise

  Serveur (messagerie, DNS, …) indisponible, vol d'information (liste des clients, sources d'un logiciel, secrets de fabrication, …), destruction de données *(hacker* chiffre des données sensibles et demande une rançon*),* copie illicite d'impressions, modification de données (mon salaire), usage de faux (je me fais passer pour le CEO), …

  …, image de marque, confiance des clients, …

- Qui doit assumer les risques ?

Computer Security Institute : Report 2010-2011

Type of Attack
Malware infection
Bots / zombies within the organization
Being fraudulently represented as sender of phishing messages
Password sniffing
Financial fraud
Denial of service
Extortion or blackmail associated with threat of attack or release of stolen data
Web site defacement
Other exploit of public-facing Web site
Exploit of wireless network
Exploit of DNS server
Exploit of client Web browser
Exploit of user's social network profile
Instant messaging abuse
Insider abuse of Internet access or e-mail (i.e. pornography, pirated software, etc.)
Unauthorized access or privilege escalation by insider
System penetration by outsider
Laptop or mobile hardware theft or loss
Theft of or unauthorized access to PII or PHI due to mobile device theft/loss
Theft of or unauthorized access to intellectual property due to mobile device theft/loss
Theft of or unauthorized access to PII or PHI due to all other causes
Theft of or unauthorized access to intellectual property due to all other causes

Types d'attaques

- Grande variétés de techniques

  virus, ver, *buffer overflow*, code mobile, *backdoor*, cheval de Troie, *rootkit*, *spyware*, *phishing*, *javascript*, injection SQL, *Cross Site Scripting*, *botnet*, *spam*, *Denial of Service*, prise de contrôle à distance, *zero day*, …

- Réelle complexité à développer une nouvelle technique d'attaque → niveau expert

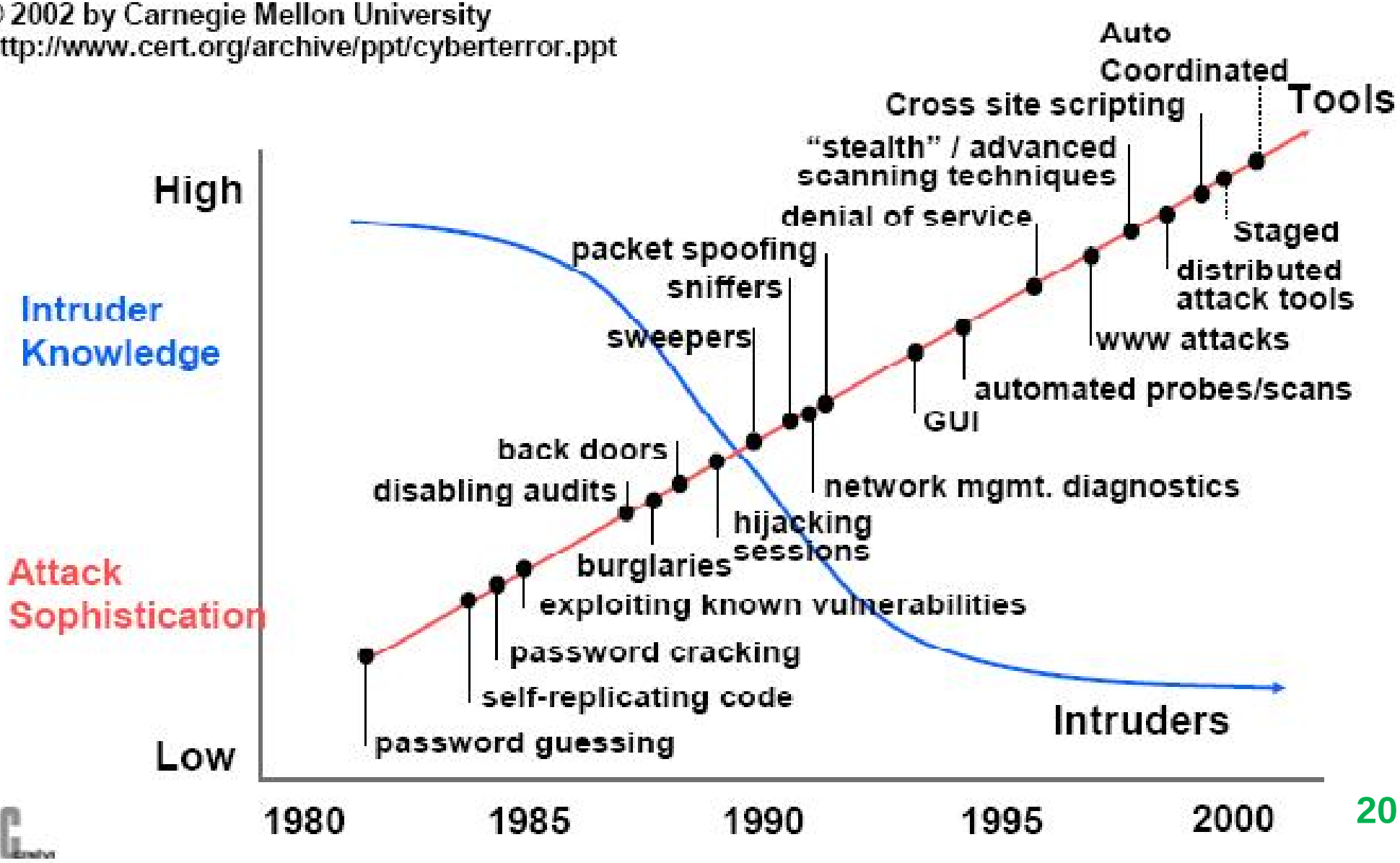
- Grande facilité à utiliser des outils puissants disponibles

- Profiter de la naïveté de sa victime → *social engineering* = attaque non technique : vous avez gagné 1 Mio CHF; il suffit de verser …

Attack Sophistication vs. Intruder Technical Knowledge

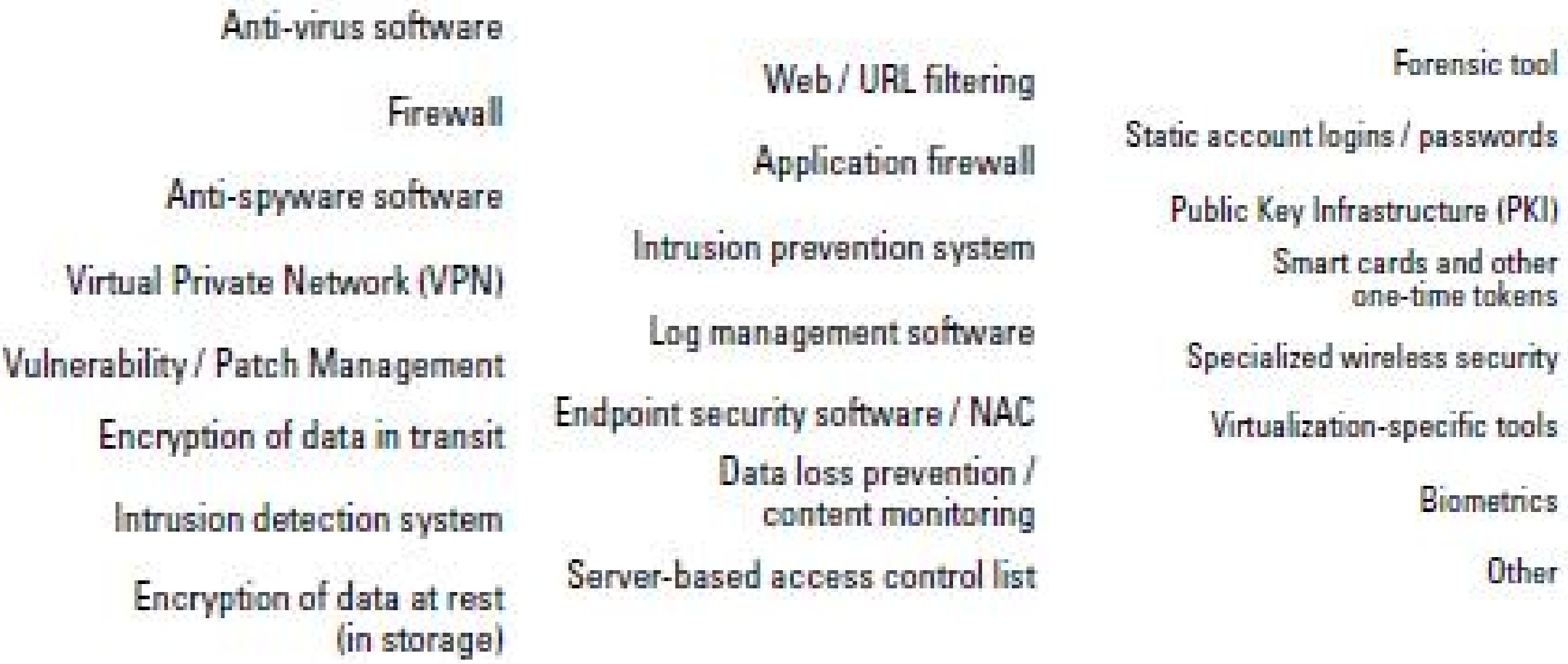
Stuxnet

http://www.cert.org/archive/ppt/cyberterror.ppt



2011

Types of Security Technology Used (CSI Report)



- Mode de fonctionnement ?
- Liste d'exclusion (antivirus), de filtrage (URL) → **black list**
- Tout ce qui n'est pas écrit est interdit (militaire SNLE) → **white list**

Sous-marins Nucléaire Lanceur d'Engins

Principaux risques (SecureWave)

Applications

Autorisées

Système d'exploitation

Logiciels métiers

Non-Autorisées

Jeu, Shareware

Logiciel piraté

Logiciels que l'utilisateur ne doit pas accéder

Malware

Connus

Virus

Ver

Cheval de Troie

...

Inconnus

Zero day exploit

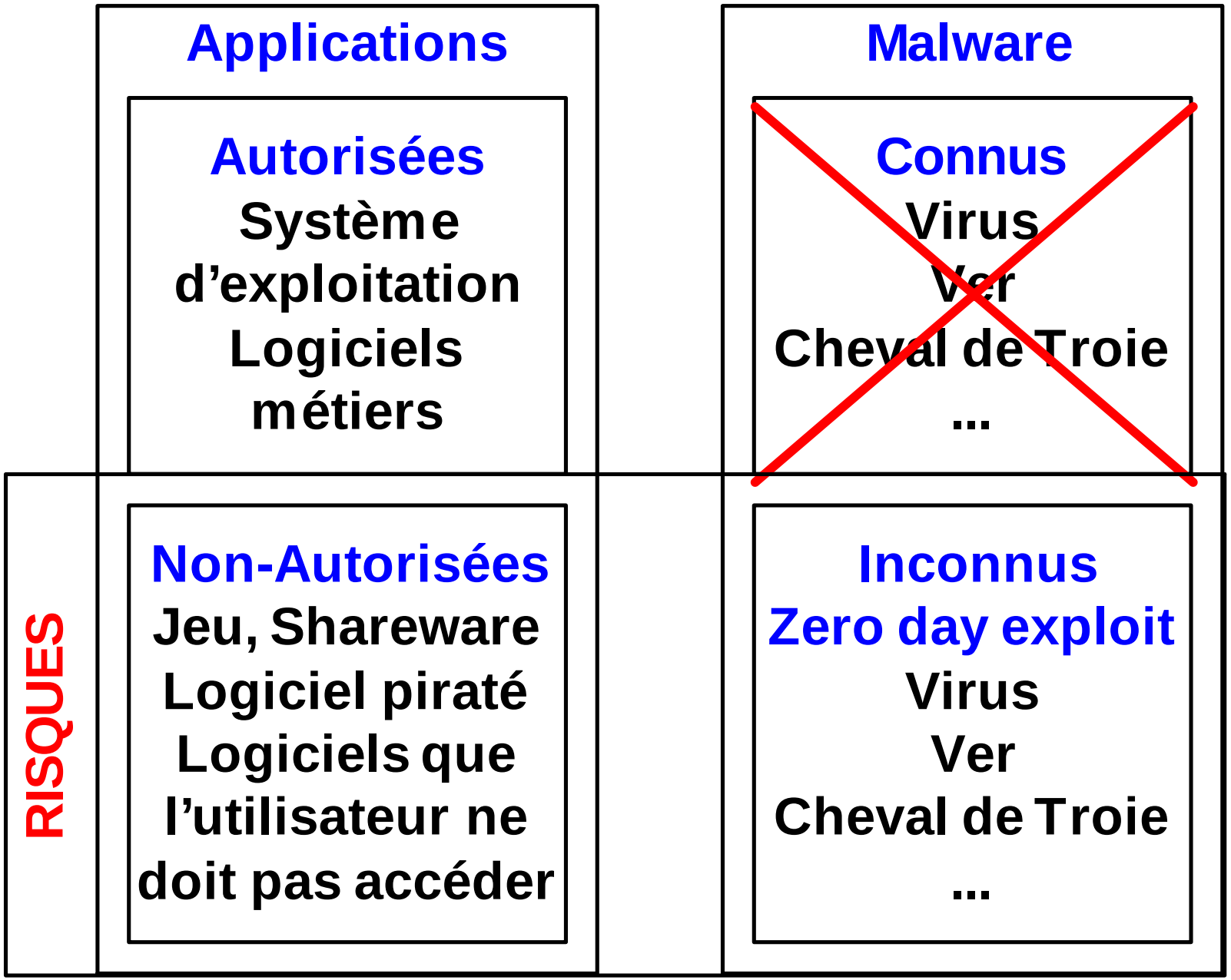
Virus

Ver

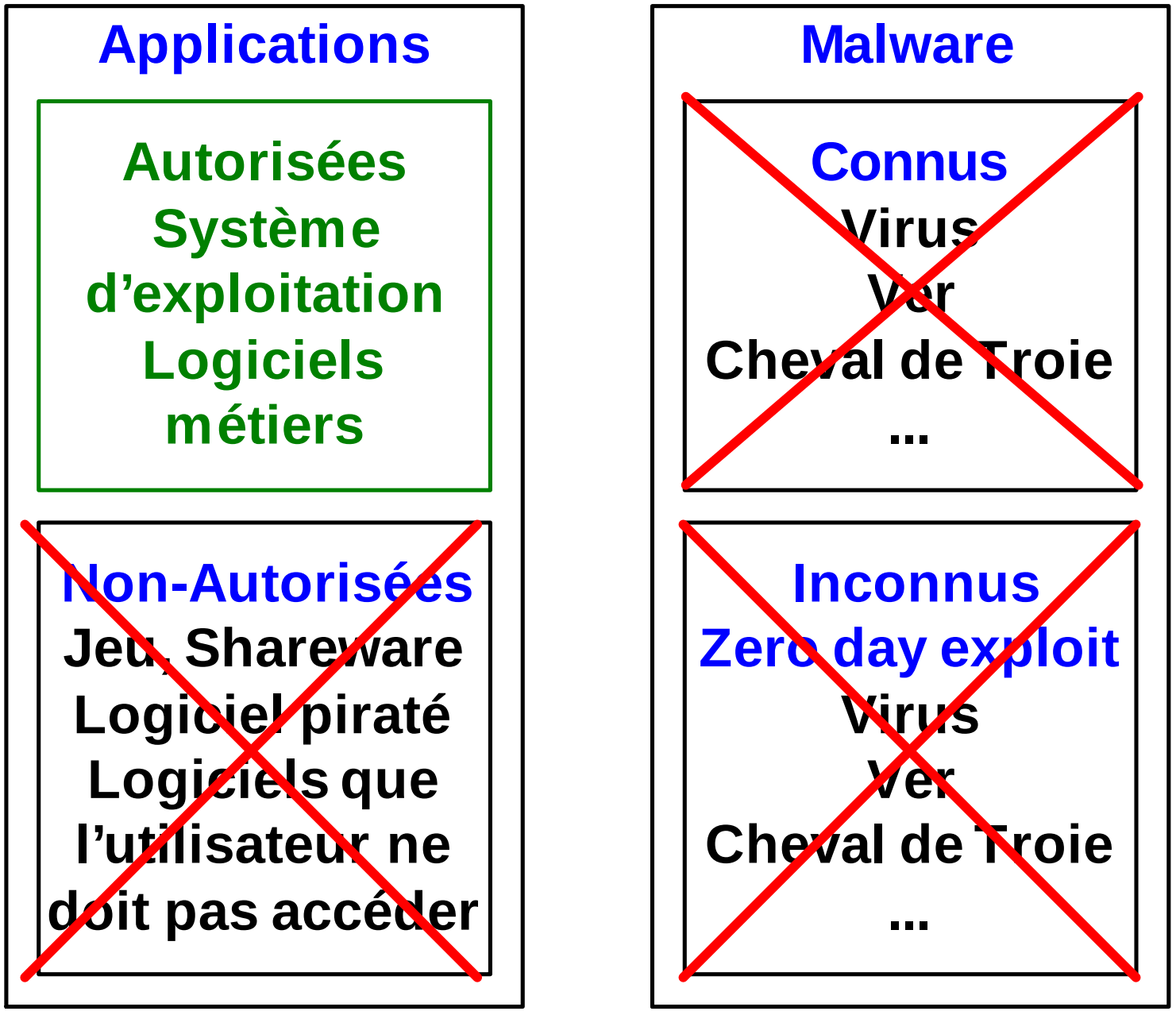
Cheval de Troie

...

Protection de type black list

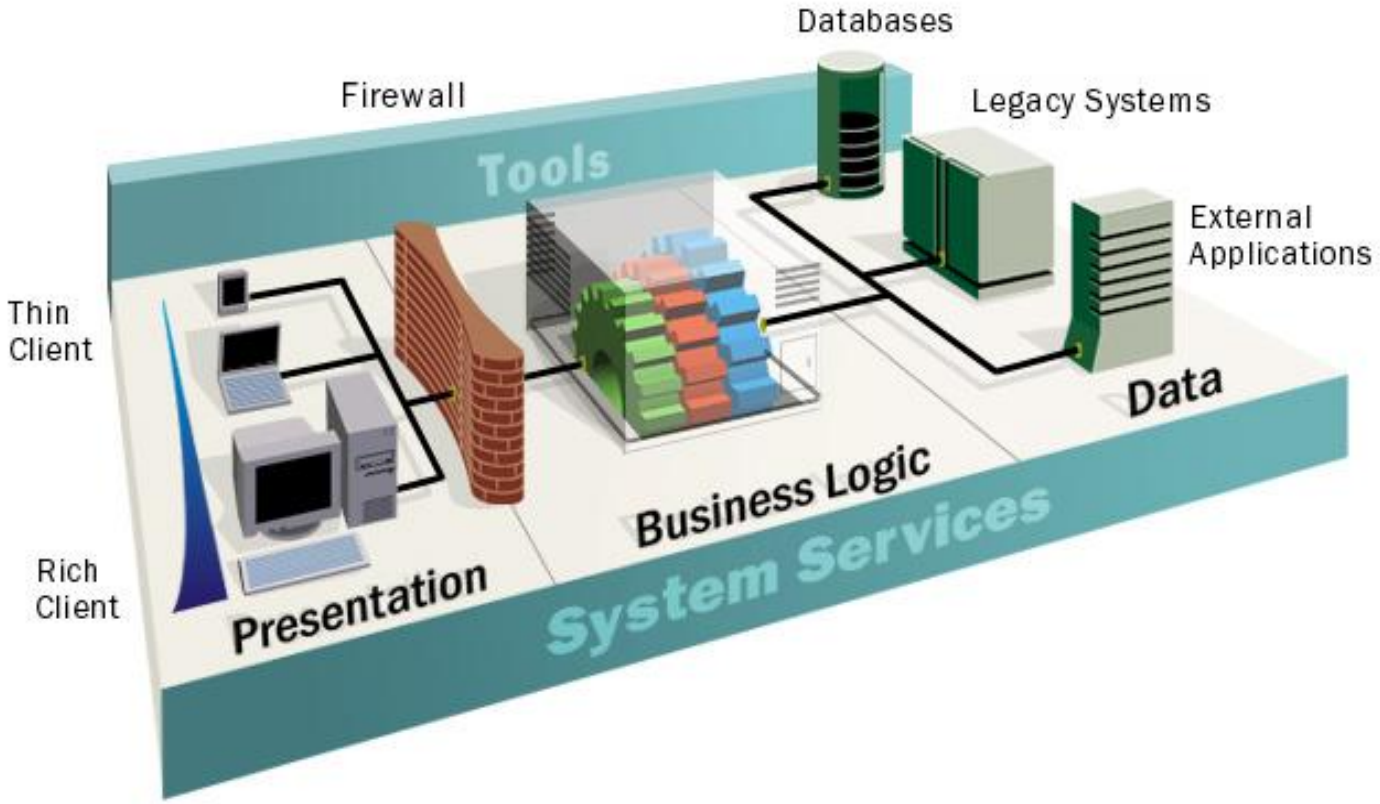


Protection de type white list



Composants d'un SI

- Personnes
- Données
- Logiciels
- Matériels
- Réseaux



→ Risques à tous les niveaux !

Attributs de la qualité de l'information

L'information doit être

- Contenu — exacte, exhaustive, concise, …
- Forme — claire, structurée, …
  Graphic User Interface (GUI) versus
  Command Line Interface (CLI)
- Temps — actuelle, à jour, …
  durée de vie

→ En tenir compte au quotidien et professionnellement

Classification des 6 principes de défense (1)

- Authentification

  Reconnaître de façon sure une personne, un serveur, une application, partenaire de confiance, …

  Risque : usurpation d'identité, vol d'identité, …

- Intégrité

  Etre certain que les données n'ont pas été modifiées

  Risque : modification des données, …

- Confidentialité

  Restreindre l'accès des données aux personnes autorisées

  Risque : vol d'information, …

Classification des 6 principes de défense (2)

- Disponibilité

  Garantir le fonctionnement d'un service → redondance, *backup*

  Risque : indisponibilité d'un serveur, perte de données, déni de service

- Autorisation

  Donner les droits *(Read/Write)* appropriés → rôles

  Risque : modification / vol des données

- Auditabilité (traçabilité)

  Permettre de retrouver les indices (preuves) d'un acte (délit)

  Risque : effacement des logs du serveur par le *hacker*, …

Politique de sécurité du SI (POSI)

- Evaluation des risques

  Quels sont les impacts (financiers, …) si le serveur de messagerie de notre société est indisponible pendant 24 heures ?

- Ne pas sous-estimer l'aspect humain

  L'homme est la source de nombreux problèmes !

- Contrôle d'accès physique (bâtiment, porte, serrure, …)

  Avez-vous fermé à clé la porte de votre voiture ?

- Tenir compte du cadre légal (LPD = Loi sur la Protection des Données, ...), réglementaire et normatif (ISO, CobiT, …)

- Plan catastrophe (en cas de grippe AH1N1), …

Analyse des risques

Axer l'appréciation de la sécurité sur l'évaluation des risques

- Quantifier les risques → estimer le préjudice financier
- Diminuer les risques
- Installer des mesures contres les risques importants *(backup, …)*
- Se protéger contre les risques non-couverts (assurances, …)

Utiliser une norme

- ISO 17799 *Code of practice for information security management – British Standards Institution*
- CobiT *Control objectives for information & related Technology*

Tableau d'analyse du risque

		Probabilité		
		Faible	Moyen	Important
Impact	Faible	5	4	3
	Moyen	4	3	2
	Important	3	2	1

diminuer l'impact

diminuer la probabilité

5 → risque acceptable

1 → risque critique

Introduction à la théorie de l'information (rappel)

- Codage de canal

  Ethernet, spectres NRZ et biphase, *half-duplex & full duplex*

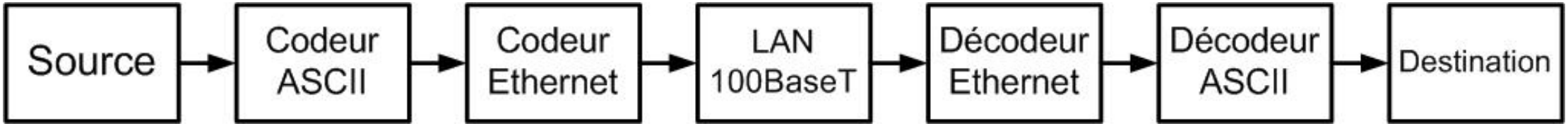
- Codage de source

  ASCII, quantité de décision, code de longueur variable, entropie, codage de texte

Système de communication

- **Modèle théorique**

- **Illustration**



- **Distinguer entre codage-décodage de source et de canal**

- **Quelles sont les contraintes liées au canal ?**

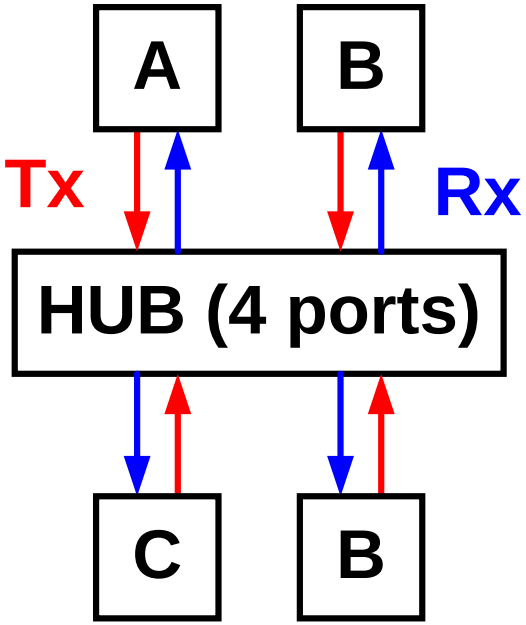
Ethernet 10 Base T (3)

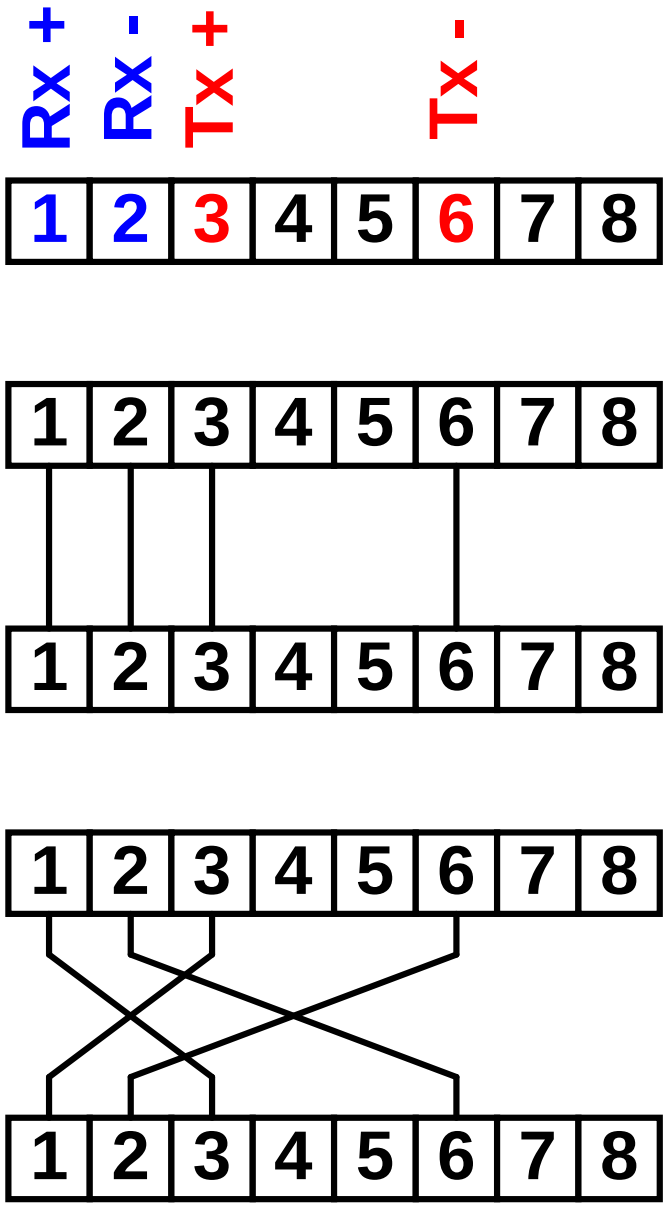
- Transmission sérielle

  → câblage simplifié

  2 paires torsadées

  distance max = 100 m





Straight cable

Crossover cable

Contraintes liées au canal

- Transmission sérielle pour simplifier le câblage (économie)

- Tout (n'importe quel) canal est limité en fréquence

  Effet passe-bas de la paire symétrique

  Plusieurs canaux multiplexés en fréquence

- Bande passante définit l'intervalle de fréquence utilisable

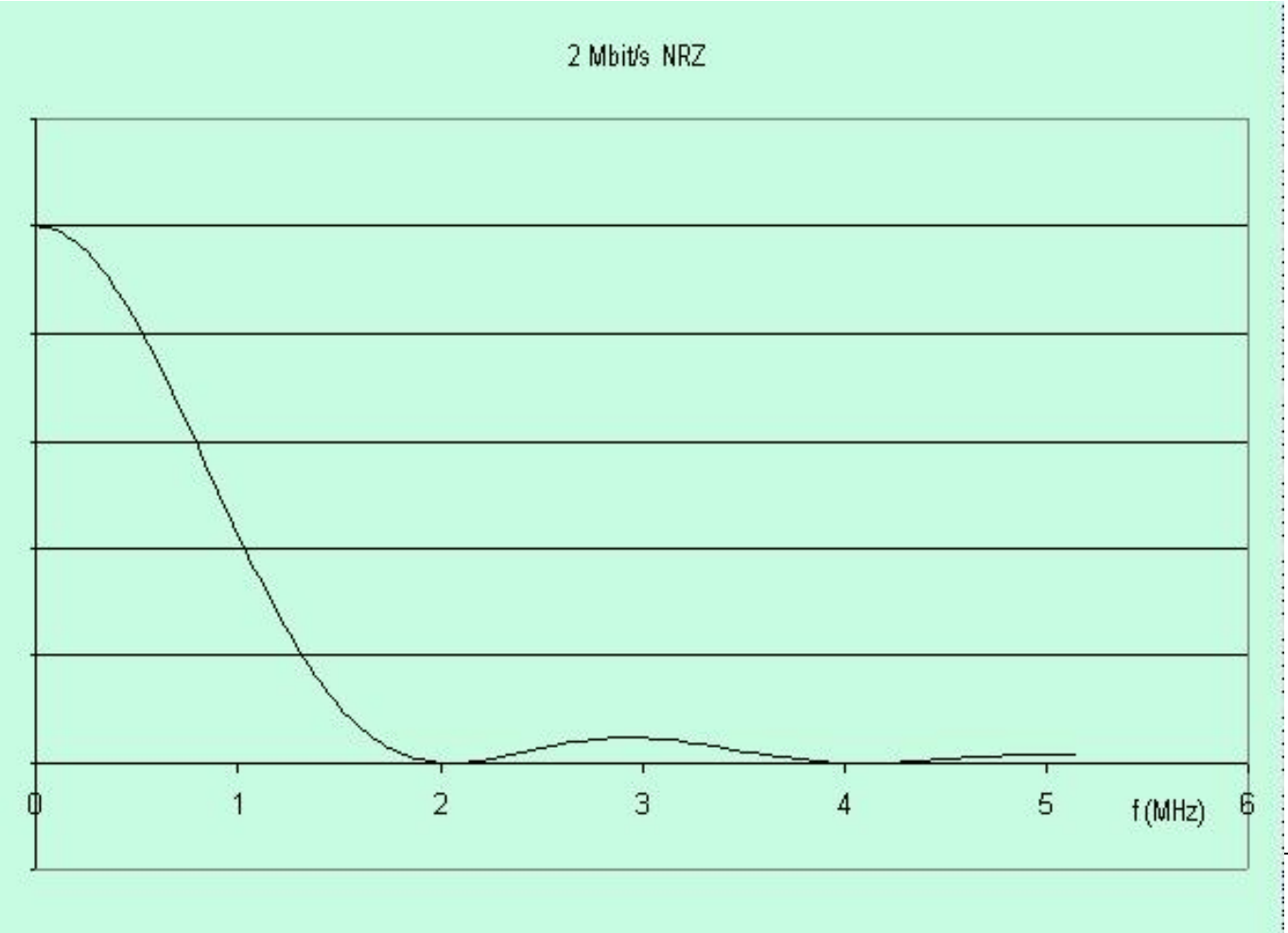
  Distinguer entre bande passante garantie (3,1 kHz en téléphonie analogique) et bande passante disponible (souvent supérieure)

- Spectre d'un signal électrique = composantes fréquentielles du signal

Codes NRZ

- Etat logique 0 codé avec une tension U0
- Etat logique 1 codé avec une tension U1
- Codage NRZ (Non Return to Zero)

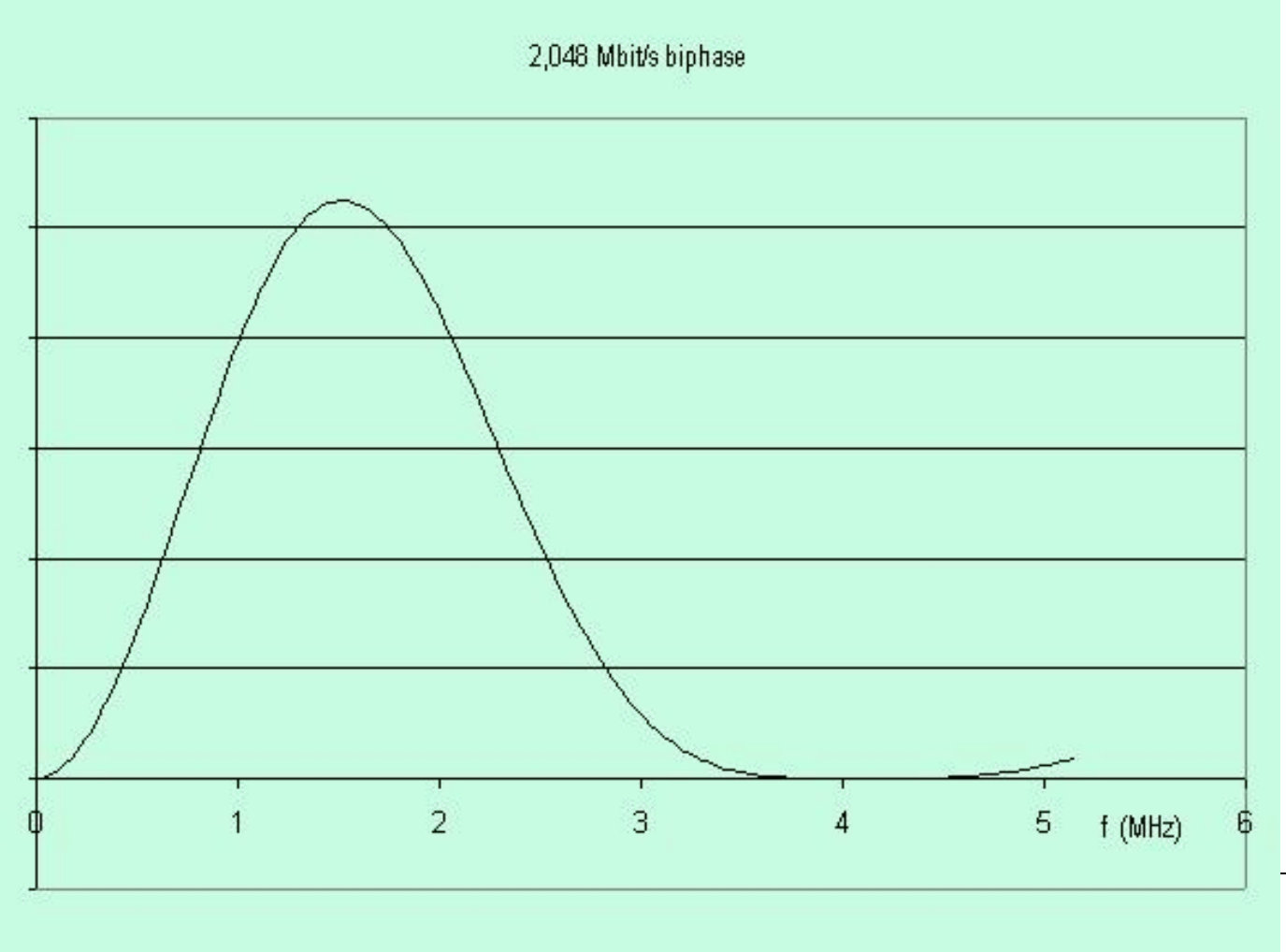
- Spectre



Codes biphase

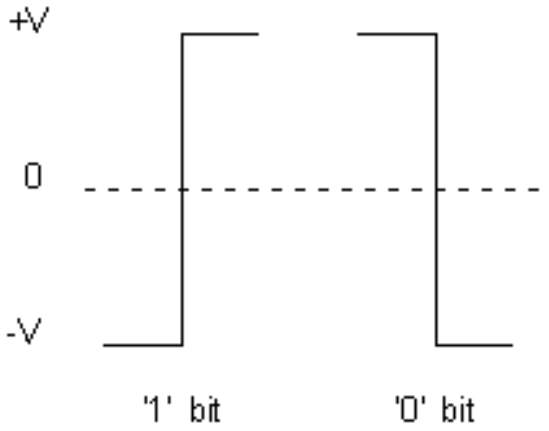
- Etat logique 0 codé avec un flanc descendant
- Etat logique 1 codé avec un flanc montant
- Autres variantes possibles : code Manchester, …

- Spectre

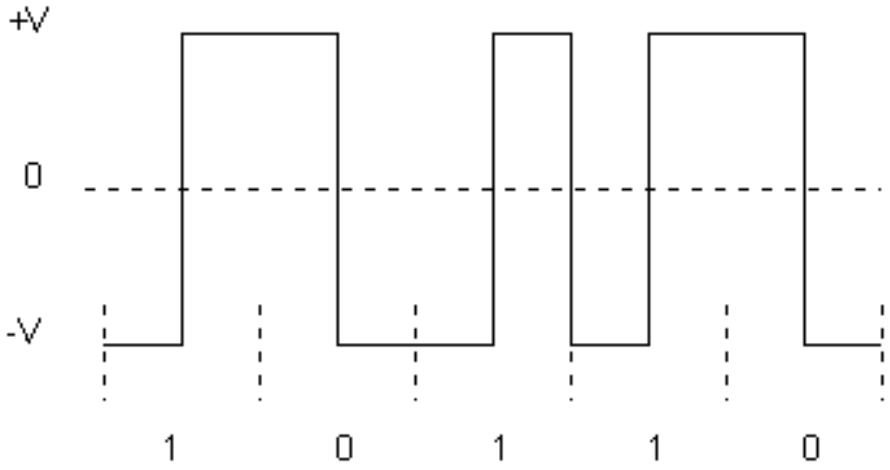


Ethernet 10BaseT

- Débit binaire de 10 Mbit/s
- Transmission en bande de base par code Manchester
- Spectre biphase



	volt
+V	0.85
idle	0
-V	-0.85



Problématique liée au canal

- Transmission limitée par le débit binaire (exemple LAN 1 Gigabit/s)

- Récupérer l'horloge (transmissions synchrones comme Ethernet) pour pouvoir lire le bloc reçu

- Mode de transmission

  Simplex, *full duplex*, *half duplex*

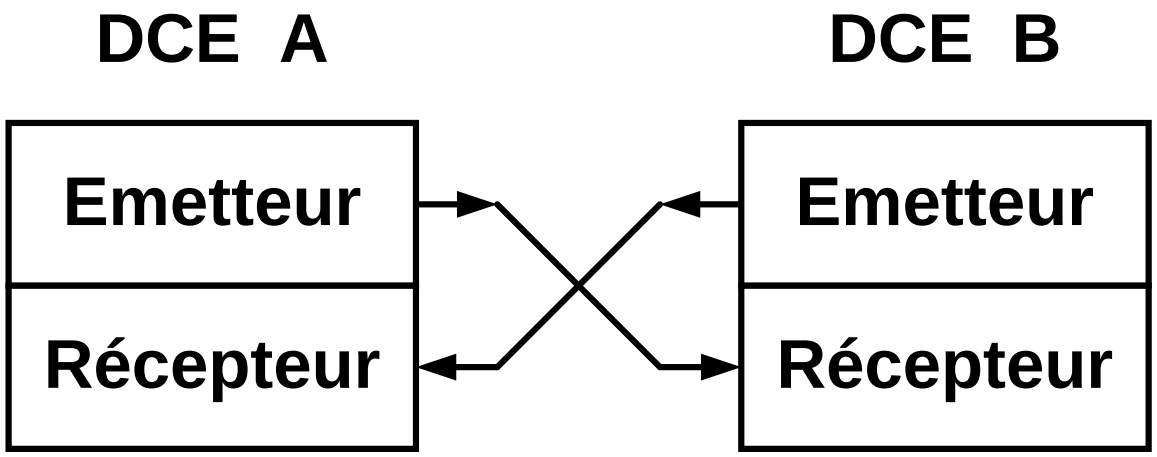
  Exemple 1 : *hub* Ethernet 10BaseT n'autorisant que le mode half-duplex

  Exemple 2 : *switch* Ethernet autorisant le mode *full-duplex*

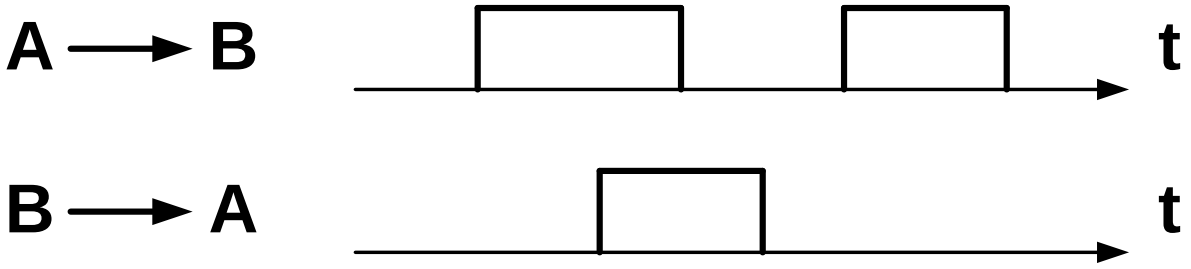
  Démo *View Status* → possibilités de l'interface Ethernet

Duplex intégral (full duplex)

- Dans le mode duplex intégral, chaque sens dispose d'un canal de transmission :

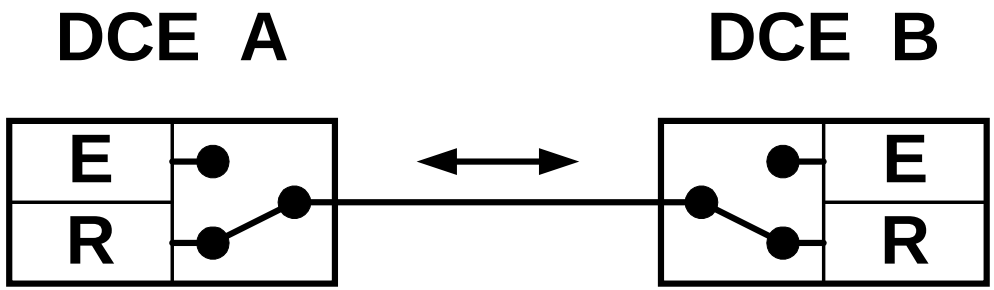


- Les 2 extrémités A et B peuvent donc émettre et recevoir simultanément

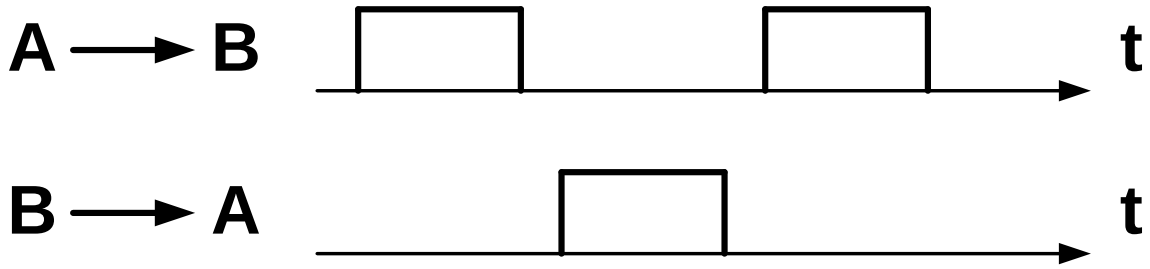


Semi-duplex (half duplex)

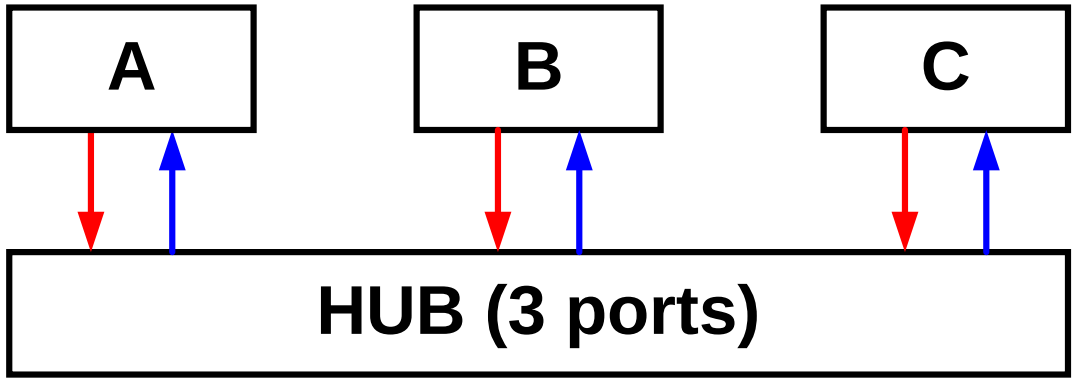
- En semi-duplex , un seul canal est partagé entre les 2 sens de transmission :



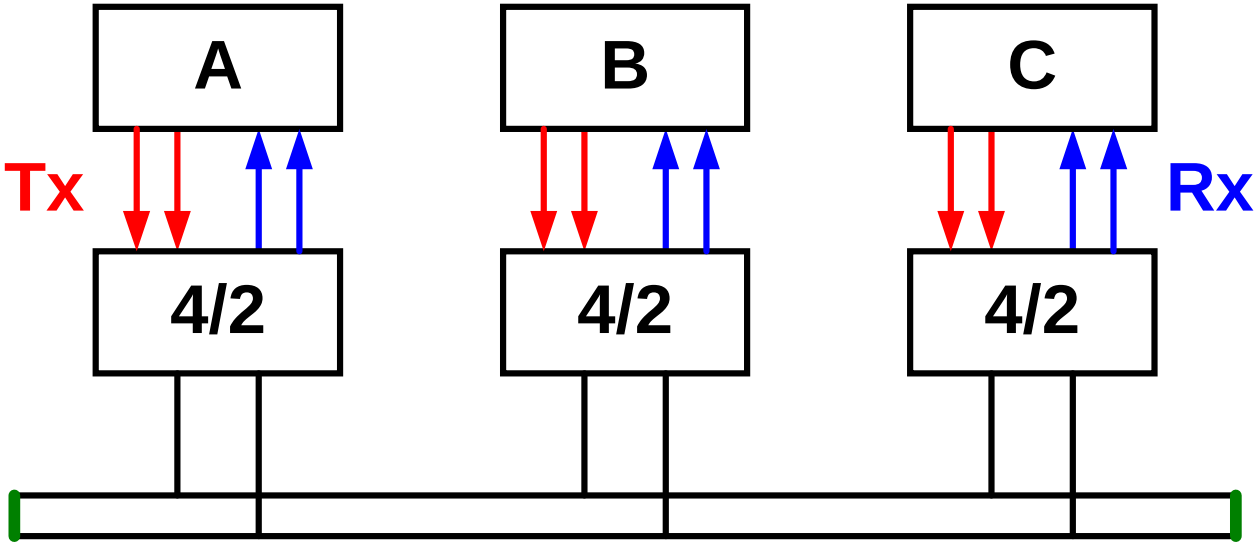
- Ce mode requiert une discipline de dialogue (protocole)
- Chaque poste est normalement en état de réception; état qu'il quitte pour émettre (principe du *talky-walky*)



Ethernet 10 Base T fonctionne en mode half duplex



- Schéma bloc du *hub*



Conversion 4 fils / 2 fils

Terminaisons (50 Ohm)

Types de source

- Source continue

  Délivre un signal qui contient une infinité de valeurs instantanées

  Grandeur analogique

  Exemple : microphone, générateur sinusoïdal, capteur CCD (caméra), …

  Domaine d'application = *Voice & Video*

- Source discrète

  Produit un ensemble fini de symboles

  Grandeur numérique codée sur N bits

  Exemple clavier, lecteur code-barres, …

- Conversion A/N & N/A

Codage de l'information (1)

- Les données numériques sont produites par une source discrète (par exemple un clavier), qui produit de l'information à partir d'un nombre fini N de caractères (lettres, chiffres, signes,...)

- Chacun de ces caractères est désigné par une expression logique codée, correspondant généralement à une seule combinaison binaire

- Le tableau de correspondance bilatérale entre les N caractères et les combinaisons binaires constitue le code

- Il permet les opérations de codage et décodage

Codage de l'information (2)

Les différents codes se distinguent par :

- leur richesse : majuscules, minuscules, signes spéciaux
- leur efficacité au sens de l'économie de configurations binaires qu'ils requièrent pour représenter un caractère

Les critères de décision permettant le choix d'un code dépendent de l'application :

- compatibilité avec d'autres équipements (norme)
- nature des informations à échanger (texte, code binaire, musique, …, données confidentielles)

Code ASCII (1)

- *American Standard Code for Information Interchange*
- Code 7 bits → nombre de combinaisons = ?
- Tableau ASCII → http://www.asciitable.com/

Dec	Hex	Code	CTRL
49	31	1	
50	32	2	
13	0D	CR	<CTRL><M>
17	11	DC1	<CTRL><Q> = Xon
19	13	DC3	<CTRL><S> = Xoff

Code ASCII (2)

NUL	*Null character*	DLE	*Data Link Escape*
SOH	*Start of Header*	DC1	*Dev. Cont. 1* (Xon)
STX	*Start of Text*	DC2	*Device Control 2*
ETX	*End of Text*	DC3	*Dev. Cont. 3* (Xoff)
EOT	*End of Transmission*	DC4	*Device Control 4*
ENQ	*Enquire*	NAK	*Neg. Acknowledgem.*
ACK	*Acknowledge*	SYN	*Synchronize*
BEL	*Bell*	ETB	*End of Text Block*
BS	*Bachspace*	CAN	*Cancel*
TAB	*Tab*	EM	*End of Media*
LF	*Line Feed*	SUB	*Substitute*
VT	*Vertical Tab*	ESC	*Escape*
FF	*Form Feed*	FS	*Form Separator*
CR	*Carriage Return*	GS	*Group Separator*
SO	*Shift Out*	RS	*Record Separator*
SI	*Shift In*	US	*Unit Separator*

Code ASCII (3)

- Sur un clavier (environ 60 touches), les 128 combinaisons sont générées de la façon suivante :

Colonne	00-1F	20-3F	40-5F	60-7F
	caractères de contrôle	chiffres signes	lettres maj.	lettres min.
			← SHIFT ←	
	←	CONTROL ←		←

- Les 2 premières colonnes (0-1) sont affectées à des caractères de contrôle *(control character)* pas imprimables

- Ex1 : Quels caractères sont interprétés par cmd.exe ?

Quantité de décision

- Nombre de bits n nécessaires pour coder l'information contenue dans un message composé de N caractères

$$n = -\log_2 (1/N) = -(\log_{10} (1/N) / \log_{10} 2)$$

- Nombre de combinaisons N possible d'un code de n bit

$$N = 2^n$$

$$= 2, 4, 8, 16, 32, 64, \ldots \text{(puissances de 2)}$$

- Déterminer la quantité de décision du message αγβαβαδα

  4 caractères différents → 2 bit

- Solution simpliste

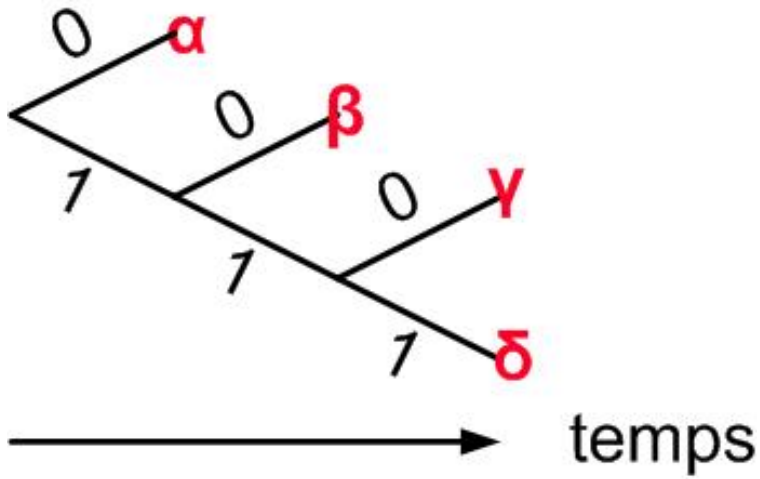
Code de longueur variable

- Il peut sembler judicieux d'utiliser un code de longueur variable si la fréquence des symboles du message varie → Voir code Morse

- Fréquence des symboles = probabilité p des symboles =
  p(α) = 0.5, p(β) = 0.25, p(γ)=p(δ)=0.125

- Principe de codage : donner au caractère le plus fréquent la combinaison binaire la plus courte

α	**0**
β	**10**
γ	**110**
δ	**111**

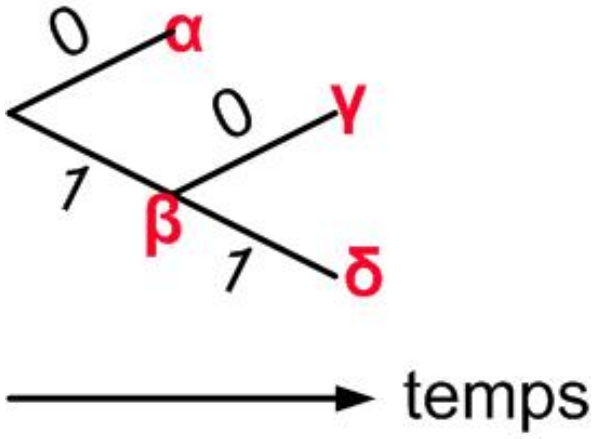


Code de longueur variable (suite)

- Déterminer le nombre de bits par symbole du message αγβαβαδα

  Message = 01101001001110 de 14 bit → 14/8 = 1.75 bit

- Le code proposé est dit séparable car il permet au récepteur de retrouver le message contrairement à une solution qui utiliserait les nœuds de l'arbre



Entropie

- Entropie H = valeur moyenne de la quantité d'information Hi portée par chaque symbole du message

  $H = \Sigma\ p(i)\ H_i = -\ \Sigma\ p(i)\ \log_2 [p(i)]$ en bit

- Calculer l'entropie du message αγβαβαδα

  $H = -\ \Sigma\ p(i)\ \log_2 [p(i)] = -\ 0.5\ \log_2\ 0.5$
  $-\ 0.25\ \log_2\ 0.25$
  $-\ 0.125\ \log_2\ 0.125$
  $-\ 0.125\ \log_2\ 0.125$
  $= 1.75$ bit

- L'entropie permet donc de calculer le nombre minimum de bit nécessaire pour coder un message

Entropie (suite)

- Quelle est la condition pour que l'entropie d'un message soit maximum ?

  Tous les symboles sont équiprobables

  $H = - \Sigma\ p(i)\ \log_2 [p(i)]$

  $\rightarrow H_{MAX} = - \Sigma\ 1/N\ \log_2 [1/N] = \log_2 N$

  Avec 4 symboles ($p(i) = 1/N = 0.25$) $\rightarrow H_{MAX} = \log_2 N =$ 2 bit

- Voir aussi http://en.wikipedia.org/wiki/Huffman_coding

Illustration avec Notepad

- Données = abc
- Save As ANSI (ASCII) **61 62 63**
- Save As UTF-8 **EF BB BF 61 62 63**
- Save As Unicode **FF FE 61 00 62 00 63 00**
- Save As Unicode Big End **FE FF 00 61 00 62 00 63**
- UTF (Universal char set Transformation Format) → 1,2,3 ou 4 byte
  http://fr.wikipedia.org/wiki/UTF-8

Ex 2

- Quel est l'outil capable d'afficher les données brutes précédentes ?
- http://notepad-plus-plus.org/
- http://winhex.com/winhex/index-f.html