

MINISTÈRE DE L'ÉDUCATION NATIONALE, DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE





# > SCIENCES ET TECHNOLOGIE

Approfondir ses connaissances

Matériaux et objets techniques

# Stockage des données

### Introduction

Les fiches connaissances ont pour objectif de présenter les principales connaissances et savoir-faire scientifiques et technologiques du thème « Matériaux et objets techniques » du programme de « Sciences et technologie » du cycle 3.

Ces fiches ne constituent en aucune manière un manuel d'enseignement des sciences et technologie ni un document pédagogique qui décrirait des situations d'enseignements. Elles sont destinées aux professeurs, afin de les aider à maitriser les concepts et notions disciplinaires. Elles ne sont pas destinées aux élèves. Les contenus peuvent aller au-delà de ce qui est attendus dans les programmes.

## Références au programme

### REPÉRER ET COMPRENDRE LA COMMUNICATION ET LA GESTION DE L'INFORMATION

- Environnement numérique de travail.
- Le stockage des données, notions d'algorithmes, les objets programmables.
- Usage des moyens numériques dans un réseau.
- Usage de logiciels usuels.

Les élèves apprennent à connaître l'organisation d'un environnement numérique. Ils décrivent un système technique par ses composants et leurs relations. Les élèves découvrent l'algorithme en utilisant des logiciels d'applications visuelles et ludiques. Ils exploitent les moyens informatiques en pratiquant le travail collaboratif. Les élèves maitrisent le fonctionnement de logiciels usuels et s'approprient leur fonctionnement.

### Les dossiers

Les dossiers sont des « classeurs » dans lesquels il est possible de déposer d'autres dossiers (appelés sous-dossiers) ou bien des fichiers.

L'arborescence permet d'organiser ses dossiers et fichiers, de manière rapide et efficace.

Cela permet d'organiser les supports de stockage pour retrouver des fichiers. Sa forme rappelle celle d'un arbre (on part de la racine, le tronc relit l'ensemble des branches constituées d'une ou plusieurs feuilles).



Retrouvez Éduscol sur









Exemple de dossier :

Le dossier TSI contient 17543 dossiers et 192 785 fichiers. La taille sur le disque est de 118 Go.

### **Vocabulaire**

Un « Bit » (contraction américaine de Binary digiT) est un digit du système binaire (valeur 0 ou 1).

Un « Mot » (Word en américain) est un ensemble de bits dont il faut préciser le nombre

Un « Octet » est un mot de 8 bits. On dit aussi « Byte »

### Les Fichiers

Les fichiers sont des documents numériques sous forme de textes, tableau numériques, images, vidéo, etc., qui sont créés et/ou rangés dans des dossiers, comme sont classées les documents papiers dans un dossier cartonné. Un fichier se distingue d'un dossier par une petite icône qui le caractérise.

Il est possible de distinguer la nature des fichiers grâce à leur extension. Cette dernière est composée d'un point et de 3 ou 4 caractères (lettres, chiffres, ...) qui terminent le nom du fichier. Par exemple, un fichier « lettre.odt » est un document texte. Cela donne aussi l'indication pour le logiciel qui pourra l'ouvrir.

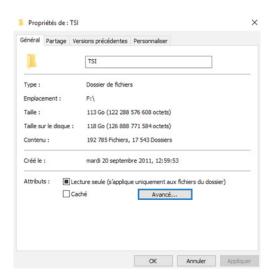
Le format de fichier signalé par son extension (Doc, PDF, JPEG, AVI, mpeg, etc...) est la convention selon laquelle les informations sont numérisées et séquencées dans le fichier.

### La taille d'un fichier s'exprime en multiples de nombre d'octets

eduscol.education.fr/ressources-2016 - Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche - Mars 2016

Les puissances de 10 sont utilisées pour classer en poids croissants le nombres de bits ou d'octets contenus dans un fichier

- 1 Kilo bits = 103 bits; 1 kilo octets = 103 octets, on note 1 kilo octets, 1ko
- 1 Méga bits = 106 bits; 1 Mega octets = 106 octets, on note 1 méga octets, 1Mo
- 1 Giga bits = 109 bits; 1 Giga octets = 109 octets,
- 1 Tera bits = 1012 bits; 1 Tera octets = 1012 octets on note 1 Tera octets, 1To











#### Remarque:

Depuis décembre 1998, une norme internationale a imposé que les préfix kilo, méga, giga, téra, etc; correspondent aux mêmes multiplieurs que dans tous les autres domaines (avant 1998 1ko= 1024 octets soit 210 octets).

Sur l'illustration ci-dessus, il est intéressant de noter que Microsoft, même dans sa version 10 de Windows, utilise encore la notation antérieure à 1998, sans respecter la norme.

$$126\ 888\ 771\ 584\ octets = \frac{126\ 888\ 771\ 584}{1024\ *\ 1024\ *\ 1024}\ Go\ soit\ 118\ Go\ (\ arrondi\ \grave{a}\ l'unit\acute{e})$$

Exemple pour une image au format bitmap (BMP):

chaque pixel en commençant par celui en bas à gauche est défini par un groupe de 3 octets représentant respectivement l'intensité en Bleu Vert Rouge.

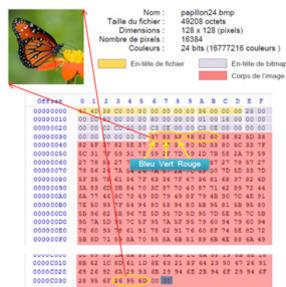
Un octet pour l'intensité du bleu (00 à FF)

Un octet pour l'intensité du vert (00 à FF)

Un octet pour l'intensité du rouge (00 à FF)

Les couleurs sont donc codés sur 3\*8 = 24 bits, soit la possibilité de coder 224 = 16777216 couleurs (plus de 16 millions de couleurs). Chaque ligne doit comporter un nombre d'octets multiple de 4 d'où les deux octets 00 et 00 à la dernière ligne.





L'image est composée de 16 384 pixels. Pour chaque pixel il faut 3 octets, un pour chacune des couleurs rouge, vert et bleu. Au total il faut donc 49 152 octets pour coder les pixels. Par ailleurs, le fichier comporte des octets réservés à l'entête du fichier et de l'image. Le volume total nécessaire est de 49 208 octets soit environ 50 ko.

## La compression de donnée

La compression de donnée consiste à obtenir des fichiers plus léger, afin par exemple d'améliorer la vitesse de transfert sur internet ou limité l'espace de stockage utilisé sur un disque dur. Il existe deux principaux types de compression :

 la compression sans perte: .zip .cab .rar .ace .7z .tar .gzip... appelée « compresser », elle consiste à coder les données binaires de manière plus concise dans un fichier et permet ainsi de retrouver la totalité des informations après une procédure de décompression ;









 la compression avec perte: .jpg .gif. Concernant principalement les fichiers média (image, son, vidéo), elle consiste en une « réduction » de l'information basée sur notre propre limite humaine à percevoir ces médias. Puisque l'œil ne perçoit pas nécessairement tous les détails d'une image, il est possible de réduire la quantité de données de telle sorte que le résultat soit très ressemblant à l'original, voire identique, pour l'œil humain.

## Supports de stockage

Les supports de stockage sont différents selon l'utilisation faite. Les paramètres importants sont la « capacité » de stockage des données et la rapidité d'écriture ou de lecture des données sur le support de stockage que l'on qualifie comme étant le « temps d'accès » :

- le disque dur : stockage direct dans l'ordinateur. Grande capacité. Il faut créer des dossiers, des sous dossiers pour s'organiser et classer ses documents (jusqu'à 2 Téraoctets soit 1012 octets):
- le disque dur externe ; stockage sur une structure externe. On peut transporter ses documents. Il possède une grande capacité de stockage (jusqu'à 1 Téraoctets);
- clé USB ; stockage externe. On peut transporter ses documents n'importe où. Capacité de stockage plus faible (1 Go à 256 Go).

Pour aller plus loin, distinguons la mémoire vive (RAM) et morte (ROM). La mémoire « vive » (RAM) est volatile, c'est à dire le contenu disparaît en absence d'alimentation, contrairement à la mémoire dite « morte » (ROM) ou les données sont conservées en l'absence d'alimentation.

TYPE	DÉFINITION	CARACTÉRISTIQUES
ROM	Read Only Memory (Mémoire en lec- ture seule)	Historiquement, les ROM (Read Only Memory) étaient effectivement des mémoires en lecture seule. Leur grosse différence avec les RAM est que leur contenu perdure malgré l'absence d'alimentation. Elles seront donc par exemple très utiles pour stocker les programmes et informations de démarrage de l'ordinateur (BIOS, Setup CMOS). Mais leurs fonctions ne se limitent pas à celle-ci.
		Actuellement, on utilise le plus souvent des EEPROM (Electrically Erasable Programmable ROM).
		Le concept de ROM :
		Si l'appellation ROM désigne le plus souvent les puces mémoires (cf photo ci-contre), on ne peut oublier :
		• les CD-ROM, les DVD-ROM
		<ul> <li>les Bluray sont des ROM</li> <li>les disques durs HDD (Hard Disk Drive) et SSD (Solid State Drive) sont des ROM (car non volatiles)</li> <li>les clefs USB sont des ROM (mémoire flash)</li> </ul>
		<b>Temps d'accès</b> : pour une puce ROM du type de celles utilisées pour le BIOS ou la CMOS, le temps d'accès est de l'ordre de quelques dizaine de nanosecondes.
		Capacités : varient selon le support considéré (puce, SSD, etc.).
RAM	Random Access Memory	Les mémoires de type RAM sont des mémoires dites volatiles, c'est à dire dont le contenu disparaît en absence d'alimentation.
	(Mémoire à accès aléatoire)	Elles sont utilisées bien-sûr dans les PC et autres ordinateurs person- nels comme mémoire de travail du système. Les mémoires Cache sont également des RAM.
A. C.		Une autre grande différence avec les ROM concerne le temps d'accès des RAM. Il est beaucoup plus faible que pour la ROM. La raison principale en est une différence essentielle de structure.
		Temps d'accès : quelques nanosecondes.
		Capacité : de l'ordre de quelques Go.









## Réseaux informatiques : généralités

Un réseau informatique est un ensemble d'équipements reliés entre eux pour échanger des informations. Les solutions techniques seront très différentes en fonction de l'ordre de grandeur des distances entre équipements. Il est possible d'établir une classification selon les dimensions et portées.

DISTANCE	CATÉGORIE	
< 1 m	Réseau personnel (PAN = Personal Area Network) liaison sans fil ordinateur/souris, clavier, imprimante contrôle appareil auditif, stimulateur cardiaque	
< 1 km	LAN - Local Area Network - salle / Maison / immeuble / collège / Campus exemple : Ethernet	
< 10 km	réseau métropolitain (MAN - Metropolitan Area Network) Ville	
< 1.000 km	réseau grande distance (WAN - Wide Area Network) pays / continent	
> 1.000 km	interconnexion de réseaux grande distance (Internet) planète, interconnexion de réseaux	enannion /oro establicario no no non

## Réseau Ethernet

Un réseau (on devrait dire sous réseau car relié à un réseau plus grand) de type Ethernet comprend les équipements suivants :

• Routeur : communément appelé la box dont on dispose par le biais d'un FAI (fournisseur d'accès à internet), le routeur permet l'envoi ou la réception d'une information en dehors du réseau Ethernet (accès à Internet). La fonction du routeur est donc d'aiquiller les paquets (exemple du paquet IP présenté plus loin) recus entre les différents réseaux, c'est à dire de réaliser des passerelles entre sous-réseaux. Le routeur exerce aussi le rôle de pare-feu, il protège un réseau informatique des intrusions indésirables en filtrant les communications autorisées ou non entre deux réseaux informatiques (généralement dans un contexte domestique).

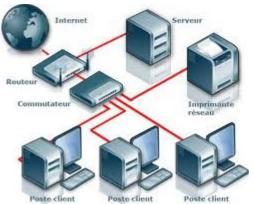
Remarque: Un pare-feu ne remplace pas un antivirus (un pare-feu ne constitue pas une protection absolue) mais l'association pare-feu avec un antivirus mis régulièrement à jour permet une protection optimale (mais toujours pas absolue).



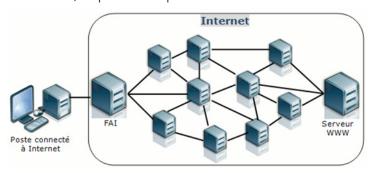




- Serveurs : ordinateurs qui fournissent des ressources partagées aux utilisateurs par un serveur de réseau.
- Clients : ordinateurs qui accèdent aux ressources partagées fournies par un serveur de réseau.
- Support de connexion : conditionne la façon dont les ordinateurs sont reliés entre eux (câble à paires torsadées et de prises RJ45, wifi ,fibre optique).
- Un commutateur (switch) qui aiguille les communications (les trames) vers la bonne destination (imprimante-serveur) en consultant dans chaque trame l'adresse matériel (adresse MAC : Médium Access Control) de l'expéditeur et du destinataire. En conservant la trace de ces adresses MAC dans sa table d'adresse, un switch est capable de transférer exactement la trame sur le port où est raccordé le destinataire.
- Périphériques partagés : imprimantes, scanner ou autres éléments utilisés par les usagers d'un réseau.



Le principe général, pour pouvoir envoyer un message d'une machine d'un réseau Ethernet à une machine d'un autre réseau Ethernet en passant par internet, consiste à passer à travers des réseaux intermédiaires, de proches en proches.



On peut illustrer ce principe en utilisant la commande «tracert» (sous Windows). L'exemple suivant montre qu'il a été nécessaire de passer par 12 réseaux intermédiaires afin d'accéder au serveur hébergeant le site www.google.fr depuis un réseau domestique ayant pour routeur la box d'adresse IP (Internet Protocol) 192.168.1.1 et le serveur google.fr a pour adresse IP 74.125.206.94.

```
sers\lionel>tracert www.google.fr
           de l'itiméraire vers www.google.fr [74.125.206.94]
m de 30 sauts :
```









## L'adresse IP (Internet Protocol)

Elle permet d'identifier chaque interface (routeur, imprimantes, carte réseau d'un ordinateur, d'un serveur, appareil connecté en Wifi comme une tablette, un robot nao, etc...) sur un réseau informatique utilisant le protocole IP (réseau Ethernet).

L'adresse IP est donc l'adresse du réseau informatique (du collège par exemple) et de la machine (dans le collège).

Une adresse IP est une adresse logique. Au contraire de l'adresse physique (MAC) qui est affectée de façon définitive à la fabrication du matériel (carte réseau) par le constructeur, l'adresse IP peut être définie et modifiée selon les besoins pour organiser des réseaux.

Remarque : l'adresse IP version 4 est utilisée encore en 2016. Mais l'explosion du nombre de machines connectées dans le monde devrait rapidement saturer le modèle actuel. Le nouveau système d'adressage (IP.v6) devrait dans les années qui viennent remplacer la version 4.

## Écriture d'une adresse IPv4

Une adresse IPv4 est codée sur 32 bits, regroupés en 4 octets.

Cela peut donner par exemple 10111111.10101000.00000000.0000001. Si on représente la valeur décimal correspondante à chaque octet on obtient 191.168.0.1.

La plus petite adresse IP possible est donc 0.0.0.0 et la plus grande 255.255.255.255 soit au total 255\*255\*255\*255 ≈ 4.2 Milliards d'adresses IP différentes. Sachant que chaque machine doit idéalement avoir une adresse IP unique (comme l'empreinte digitale d'un être humain) afin de ne pas créer d'ambiguïté entre émetteur et récepteur, et vue le nombre croissant de machine connectée au réseau internet, on voit qu'il faudra coder l'adresse IP sur davantage que 32 bits.

IPv6 propose de coder chaque machine sur 128 bits.

Une partie de l'adresse représente l'adresse du réseau (ID de réseau) et l'autre partie représente l'adresse de la machine sur le réseau (ID d'hôte).

## Masque de sous-réseau

Le masque de sous-réseau permet de savoir à quelle partie de l'adresse IP correspond l'adresse du sous-réseau (net ID en anglais) et l'adresse de la machine sur le sous-réseau (host ID).

L'adresse IP et le masque de sous réseau associé sont deux informations indissociables.

### **DÉFINITION**

Pour déterminer l'ID de réseau, un 'Et logique' est effectué entre l'adresse IP et le masque de

Pour déterminer l'ID de l'hôte, un 'Et logique' est effectué entre l'adresse IP et le 'complément à 1' du masque de sous réseau.









### Exemple:

	CODAGE DÉCIMAL	CODAGE BINAIRE
Adresse IP	192.168.1.22	11000000 10101000 00000001 00010110
Masque de sous réseau	255.255.255.0	11111111 11111111 11111111 00000000
ID de réseau	192.168.1.0	11000000 10101000 00000001 00000000
ID de l'hôte	0.0.0.22	00000000 00000000 000000000 00010110

Il est possible de retenir qu'aux bits à « 1 » du masque de sous réseau correspond l'adresse IP du sous réseau. Aux bits à « 0 » du masque de sous réseau correspond l'adresse IP de l'hôte.

Il est également possible à partir d'un réseau domestique de connaitre l'adresse IP d'un ordinateur, du routeur (passerelle par défaut) ainsi que le masque de sous réseau à partir de la commande « ipconfig » sous windows saisie sur l'invite de commandes de windows.

```
:\Users\lionel>IPconfig
Configuration IP de Windows
arte réseau sans fil Connexion au réseau local* 2 :
  Statut du média. . . . . . . . . . . . . . Média déconnecté
  Suffixe DNS propre à la connexion. . . :
arte Ethernet Ethernet :
  Suffixe DNS propre à la connexion. . . : home
  Adresse IPv6 de liaison locale. . . . .: fe80::a5f1:903b:c59e:20b0%14
  Adresse IPv4. . . . . . . . . . . . . . . . . 192.168.1.22
  Masque de sous-réseau. . . . . . . : 255.255.255.0
  Passerelle par défaut. . . . . . . : 192.168.1.1
```

Ainsi dans l'exemple ci-dessus, chaque interface du sous réseau Ethernet est identifiée par une adresse IP différente mais commençant toutes par 192.168.1.\*\*\* avec 192.168.1.0 l'identifiant d'un réseau ayant pour masque 255.255.255.0.

On peut connecter 255-1=254 machines sur un réseau ayant pour masque 255.255.255.0. Pour connecter d'avantage de machine, il faut modifier le masque en prenant par exemple 255.255.0.0. On pourra ainsi connecter 255\*254=64 770 machines.

Remarque : l'adresse IP 192.168.1.1 ne doit pas être utilisée par une machine. En effet, le routeur utilise cette adresse appelée adresse Ethernet de broadcast (diffusion) pour envoyer un message à toutes les machines du sous réseau afin de récupérer les adresses MAC de toutes les machines pour mettre à jour sa table ARP (adress resolution protocol).

De cette façon, un routeur dispose d'une table ARP permettant d'associer, à chaque adresse IP des machines d'un sous-réseau connecté, les adresses MAC correspondantes.

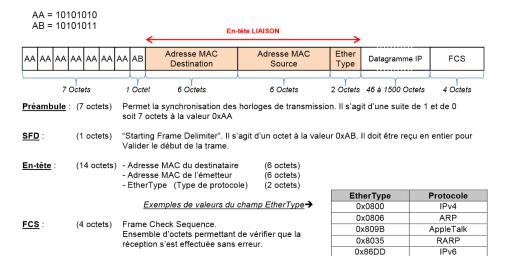




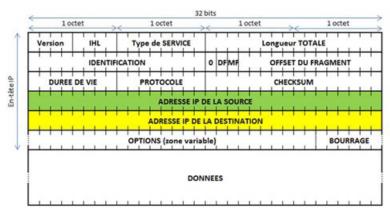


## Trame Ethernet II (version 2)

Voilà la constitution d'une trame Ethernet complète. Elle contient au plus 1526 octets.

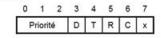


Le Datagramme IP (paquet IP) contient les adresses IP de l'émetteur et du destinataire, ainsi que l'information utile.



il indique le numéro de version du protocole IP utilisé (généralement 4). Version: (4 bits) Internet Header Lenght (Longueur d'entête). Spécifie la longueur de l'en-tête du Datagramme en nombre de mots de 32 bits. Ce champ ne peut prendre une valeur IHL: (4 bits) inférieure à 5.

(8 bits) Donne une indication sur la qualité de « service » souhaitée pour l'acheminement Type de service :



Bits 0-2	Priorité	010→Immédiate	001→Normale 000→Basse
Bit 3	D	0 = Retard standard	1 = Retard faible
Bit 4	T	0 = Débit standard 1 = Haut débit	
Bit 5	R	0 = Taux d'erreur standard 1 = Taux d'erreur faible	
Bit 6	С	0 = Coût standard 1 = Coût faible	
Bit 7	X	Réservé	

(MF - More Fragment)

0= Dernier fragment

Longueur totale :	(16 bits)	Longueur du	datagramme entier y compri	is en-tête et données m	esurée en octets.
Identification:	(16 bits)	Valeur assignée par l'émetteur pour identifier les fragments d'un même datagramme			
Flags: (3 bits) Commutate			rs de contrôle :		
(, 1000000000000000000000000000000000000		- Bit 0	Réservé, doit être laisse	éà0	
		- Bit 1	(DF - Don't fragment)	0= Fragmenté	1= Non frage

Retrouvez Éduscol sur









- Bit 2

1= Fragment

OFFSET: (13 bits) Décalage du premier octet du fragment par rapport au datagramme complet non fragmenté. Cette position est mesurée en blocs de 8 octets (64 bits).

Durée de vie : (8 bits) Temps en secondes pendant lequel le datagramme doit rester dans le réseau Si ce champ vaut 0, le datagramme doit être détruit. Ce temps diminue à chaque

passage du datagramme d'une machine à l'autre

(8 bits) Protocole: Protocole porté par le datagramme (au-dessus de la couche IP)

Valeur	Protocole		
1	ICMP		
6	TCP		
17	UDP		
Etc	etc		

Checksum: (16 bits) (Somme de contrôle) C'est une valeur qui permet de déceler une éventuelle erreur de transmission avec une très grande probabilité.

IP Source : (32 bits) Adresse IP du destinataire IP Destination: (32 bits)

(Variable) Le champ est de longueur variable. Un datagramme peut comporter 0 ou plusieurs options. Options: Le champ Bourrage n'existe que pour assurer à l'en-tête une taille totale multiple de

(Variable) Bourrage: 4 octets. Le bourrage se fait par des octets à 0.

Les informations qui doivent fondamentalement se trouver dans l'entête IP sont:

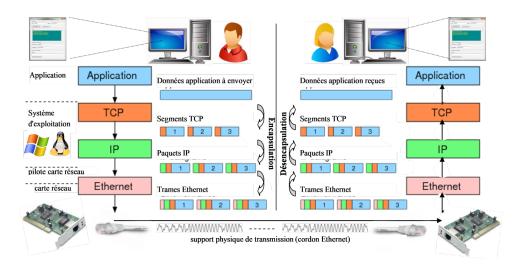
• l'adresse IP de l'émetteur :

l'adresse IP du destinataire :

le protocole utilisé sur la couche supérieure (couche transport).

## Une organisation en couche des protocoles

On peut ainsi schématiser l'implémentation par couches des protocoles de communication.



Le protocole TCP (Transmission Control Protocol - en français protocole de contrôle de transmission) est un protocole (comme UDP et bien d'autres) de transport fiable de la couche transport (comme le protocole IPv4 ou IPv6 sont des protocoles de la couche réseau).

La trame Ethernet est générée par encapsulation à partir des Données application à envoyer. L'encapsulation consiste à inclure les données d'un protocole dans un autre protocole de couche inférieure.

Les données application reçues sont obtenues par désencapsulation. La désencapsulation consiste à supprimer les entêtes renseignant des protocoles utilisés lors de l'envoi, ceci afin d'utiliser les mêmes protocoles à la réception.









### CYCLE 3 | SCIENCES ET TECHNOLOGIE | Approfondir ses connaissances

Matériaux et objets techniques

Voici deux protocoles (parmi bien d'autres) de la couche application :

- le POP (Post Office Protocol), littéralement le protocole du bureau de poste, est un protocole qui permet de récupérer les courriers électroniques situés sur un serveur de messagerie électronique;
- L' HTTP (HyperText Transfer Protocol), est un protocole de communication client-serveur développé pour le World Wide Web (accès au web).





