

Eliaz ANDRÉ LE POGAMP

eliaz.andrelepogamp@etudiant.univ-rennes1.fr

Natale ORICELLI

natale-jonas.oricelli@etudiant.univ-rennes1.fr

# Rapport de projet BDD

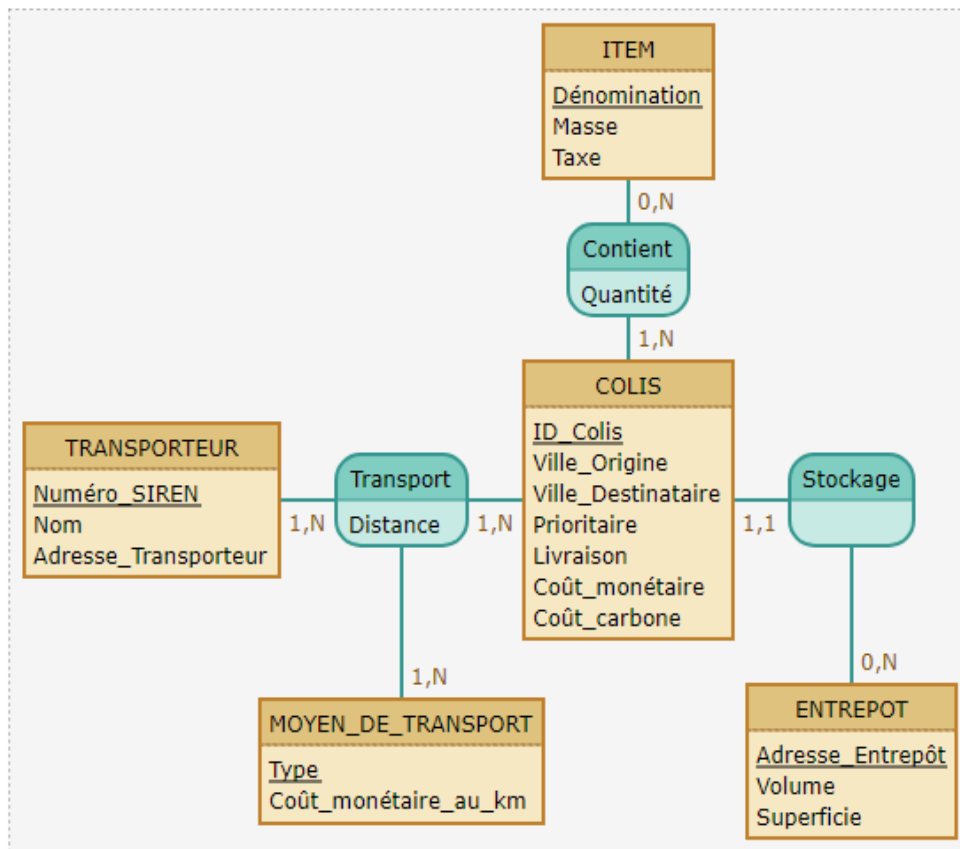
## 1.Introduction

Soyons honnêtes, l'inspiration ne nous est pas venue de suite concernant le sujet de notre base de données. Nous avons des idées, mais aucune qui puisse être mise en lien de près ou de loin avec le développement durable. En réalité, c'est en regardant un reportage sur l'impact des flux croissants de colis que nous avons décidé d'en faire notre sujet de base de données. Tout le monde a déjà commandé un colis, tout le monde en commandera à l'avenir, et si l'on en croit La Poste, 400 millions de Colissimo ont transité en France en 2022. Un nombre qui n'inclut même pas les nombreux autres services de livraison de colis tels que DHL, FedEx, MondialRelay... En prenant ces autres sociétés en compte, on pourrait estimer le nombre de colis circulant annuellement en France à au moins un milliard, une quantité colossale qui correspond à environ 15 colis par an par personne au travers des 67 millions de français. Toutes ces livraisons ont forcément un impact écologique, chaque colis devant être stocké, emballé, puis transporté, d'où le lien avec le développement durable.

Source La Poste : <https://www.lapostegroupe.com/fr/actualite/la-poste-se-prepare-a-trier-et-distribuer-plus-de-100-millions-de-colissimo-pour-noel>

## 2.Modèle conceptuel MCD

Pour notre modélisation, nous utilisons le site mocodo.net qui applique la notation MERISE. Voici le modèle :



Le modèle est composé de 5 entités et 3 relations, que l'on met en évidence par leur couleur.

Nous allons commencer par détailler les entités.

### COLIS :

- ID\_Colis\*, clé primaire, identifiant du colis, qui est un varchar.

*Note : pour simplifier, nous sommes partis du principe que tous les identifiants de colis respectaient le même format. Ce n'est pas réellement le cas, certains colis sont identifiés par des nombres entiers, d'autres par le signe « FR + [nombre\_entier] », etc*

- Ville\_Origine\*, la ville dont vient le colis, que ce soit un entrepôt ou un particulier. C'est un varchar.
- Ville\_Destinataire\*, la ville vers laquelle se dirige le colis. Un varchar également.
- Prioritaire\*, un booléen. Le colis est prioritaire si le booléen vaut 1, il ne l'est pas sinon.
- Livraison\*, un booléen. Le colis est livré à l'adresse de l'acheteur s'il vaut 1, et en point-relais s'il vaut 0.

- Coût\_monétaire\* : un double qui représente le coût de colis à l'achat en euros.
- Coût\_carbone\* : un double qui représente le coût en carbone du voyage du colis uniquement, en grammes

#### ITEM :

- Dénomination\*, clé primaire, un varchar réalisant une brève description du produit contenu dans le colis, tel que « Lance-pastèque portatif »
- Masse\*, un double, la masse d'un seul exemplaire du produit exprimée en g
- Taxe, un double qui indique le prix d'une éventuelle taxe sur le produit, s'il est importé par exemple. Il peut prendre la valeur « null » ou 0.0 si ce n'est pas le cas.

#### TRANSPORTEUR :

- Numéro\_SIREN\*, un integer clé primaire qui identifie la société du transporteur. En général, un numéro SIREN est composé de 9 entiers.
- Nom\*, un varchar qui indique le nom du transporteur, par exemple «La Poste»
- Adresse\_Transporteur, un varchar qui indique l'adresse du transporteur. Il prend rarement la valeur « null » mais c'est envisageable étant donné que certains ont des agences partout en France et pas une seule adresse unique.

#### MOYEN\_DE\_TRANSPORT :

- Type\*, clé primaire, un varchar qui peut prendre 4 valeurs selon le type de transport du colis :
  - Avion
  - Bateau
  - Ferroviaire
  - Routier
- Coût\_monétaire\_au\_km\*, un double qui indique le coût moyen d'un kilomètre parcouru par un colis en euros.

#### ENTREPOT :

- Adresse\_Entrepôt\*, clé primaire, un varchar qui renseigne l'adresse de l'entrepôt
- Volume, un entier qui renseigne le volume disponible dans l'entrepôt, en m<sup>3</sup>
- Superficie, un entier qui renseigne la surface de l'entrepôt, en m<sup>2</sup>

Maintenant on détaille les relations :

#### Transport :

- Clé primaire : Numéro\_SIREN,ID\_Colis,Type.
- Distance, un double qui indique la distance parcourue par le colis pour un certain type de transport.

### Contient :

- Clé primaire : Dénomination, ID\_Colis
- Quantité\*, un entier qui indique combien d'exemplaires du produit sont contenues dans le colis

### Stockage :

- Clé primaire : ID\_Colis, Adresse\_Entrepôt

On va maintenant justifier les **cardinalités**.

Concernant la relation **Transport** :

Un Transporteur peut proposer plusieurs moyens de transports, et il doit en proposer au minimum un. Il doit également gérer le transport d'au moins un colis et peut en transporter plusieurs → TRANSPORTEUR - 1,N - Transport

De manière assez équivalente, un Moyen De Transport doit concerner au moins un colis et peut en concerner plusieurs, tout comme il doit être utilisable par au moins un transporteur tout comme par plusieurs → MOYEN\_DE\_TRANSPORT - 1,N - Transport

Enfin, durant son trajet, un colis peut être pris en charge par un ou plusieurs transporteurs, et il peut nécessiter un comme plusieurs moyens de transport → COLIS - 1,N - Transport

Concernant la relation **Contient** :

Un colis peut contenir un seul type d'objet (qu'on appelle item) mais il en contient parfois plusieurs. Il en contient forcément au moins un. → COLIS - 1,N - Contient  
Un item peut être dans différents colis, c'est d'ailleurs le cas en général. En revanche, s'il a peu de succès ou qu'il est en rupture de stock momentanée, il est possible qu'il ne soit dans aucun colis. → ITEM - 0,N - Contient

Concernant la relation **Stockage** :

Un entrepôt doit en général contenir plusieurs colis. En théorie une grande quantité pour être utile d'ailleurs. Mais s'il est en maintenance ou la cible d'un incendie, il est possible qu'il n'en contienne aucun.  
→ ENTREPOT - 0,N - Stockage

Et pour finir, un colis une fois les produits qu'il contient emballés quitte l'entrepôt. Dans notre modélisation, on considère qu'un colis qui quitte l'entrepôt est

directement livré au destinataire sans passer par un autre entrepôt. Ainsi, chaque colis n'est stocké que par un seul et unique entrepôt. → COLIS - 1,1 - Stockage

### 3.Schéma relationnel MLD

On considère deux règles pour construire notre schéma relationnel :

- règle 1 : dans le cas où une relation fait le lien entre deux entités de cardinalité  $(0,N) / (0,N)$  ou  $(0,N) / (1,N)$  ou  $(1,N) / (1,N)$ , la relation devient une table et prend en clé primaire la fusion des clés primaires des deux entités.

- règle 2 : dans le cas où une relation fait le lien entre deux entités de cardinalité  $(0,N) / (1,1)$  ou  $(1,N) / (1,1)$ , l'entité de cardinalité  $(1,1)$  prend en attribut la clé primaire de l'entité de cardinalité  $(0,N)$  ou  $(1,N)$  qui devient donc une clé étrangère.

On commence par créer les tables qui n'ont pas de clé étrangère afin de ne pas avoir de références vers une table encore inexistante. Les clés primaires sont soulignées simple et les clés étrangères sont soulignées double.

ENTREPOT(Adresse\_Entrepôt, Volume, Superficie)

COLIS(ID\_Colis, Ville\_Origine, Ville\_Destinataire, Prioritaire, Livraison, Coût\_monétaire, Coût\_carbone, Adresse\_Entrepôt)  
Adresse\_Entrepôt → règle 2  
La clé étrangère fait référence à l'entité ENTREPOT.

ITEM(Dénomination, Masse, Taxe)

CONTIENT(ID\_Colis, Dénomination, Quantité)

La relation devient une table et prend en clé primaire la fusion des clés primaires des deux entités qu'elle lie (ID\_Colis, Dénomination) → règle 1  
Ici les clés font références aux entités COLIS et ITEM.

MOYEN\_DE\_TRANSPORT(Type, Coût\_monétaire\_au\_km)

TRANSPORTEUR(Numéro\_SIREN, Nom, Adresse\_Transporteur)

TRANSPORT(ID\_Colis, Numéro\_SIREN, Type, Distance)

À nouveau la règle 1, les clés font référence aux entités COLIS, TRANSPORTEUR et MOYEN\_DE\_TRANSPORT.

Pour que le visuel soit plus digeste, on représente les attributs **non**-obligatoires (donc ceux qui n'auraient pas eu d'astérisque \*) en fond rouge, sinon c'est lourd à regarder. De plus, cette étape est similaire à celle du 2 si vous voulez avoir des étoiles plein les yeux.

## 4.Schéma physique MPD

Voici toutes les commandes CREATE TABLE qui permettent de construire en intégralité notre base de données. Les tables sont toutes vides à la création. Nous avons bien respecté l'ordre de sorte qu'aucune clé étrangère ne fasse référence à une clé primaire d'une table pas encore créée, car cette occurrence n'est pas supporté par phpMyAdmin.

```
CREATE TABLE ITEM (  
  Denomination VARCHAR(42) PRIMARY KEY,  
  Masse        DOUBLE NOT NULL,  
  Taxe         DOUBLE  
);
```

```
CREATE TABLE ENTREPOT (  
  Adresse_Entrepot VARCHAR(42) PRIMARY KEY,  
  Volume           DOUBLE,  
  Superficie       DOUBLE  
);
```

```
CREATE TABLE TRANSPORTEUR (  
  Numero_SIREN      INTEGER PRIMARY KEY,  
  Nom               VARCHAR(42) NOT NULL,  
  Adresse_Transporteur VARCHAR(42)  
);
```

```
CREATE TABLE MOYEN_DE_TRANSPORT (  
  Type              VARCHAR(42) PRIMARY KEY,  
  Cout_monetaire_au_km DOUBLE NOT NULL  
);
```

```
CREATE TABLE COLIS (  
  ID_Colis          VARCHAR(42) PRIMARY KEY,  
  Ville_Origine     VARCHAR(42) NOT NULL,  
  Ville_Destinataire VARCHAR(42) NOT NULL,  
  Prioritaire       BOOLEAN NOT NULL,  
  Livraison         BOOLEAN NOT NULL,  
  Cout_monetaire     DOUBLE NOT NULL,  
  Cout_carbone       DOUBLE NOT NULL,  
  Adresse_Entrepot  VARCHAR(42) NOT NULL,  
  FOREIGN KEY (Adresse_Entrepot) REFERENCES ENTREPOT  
  (Adresse_Entrepot)  
);
```

```

CREATE TABLE Contient (
  ID_Colis   VARCHAR(42) NOT NULL,
  Denomination VARCHAR(42) NOT NULL,
  Quantite   INTEGER NOT NULL,
  PRIMARY KEY (ID_Colis, Denomination),
  FOREIGN KEY (ID_Colis) REFERENCES COLIS (ID_Colis),
  FOREIGN KEY (Denomination) REFERENCES ITEM (Denomination)
);

CREATE TABLE Transport (
  ID_Colis   VARCHAR(42) NOT NULL,
  Numero_SIREN INTEGER NOT NULL,
  Type       VARCHAR(42) NOT NULL,
  Distance   DOUBLE NOT NULL,
  PRIMARY KEY (ID_Colis, Numero_SIREN, Type),
  FOREIGN KEY (ID_Colis) REFERENCES COLIS (ID_Colis),
  FOREIGN KEY (Numero_SIREN) REFERENCES TRANSPORTEUR
(Numero_SIREN),
  FOREIGN KEY (Type) REFERENCES MOYEN_DE_TRANSPORT (Type)
);

```

## 5.Peuplement des tables

Après avoir créé une base vide de toute donnée, on peut maintenant la peupler avec des n-uplets.

**ENTREPOT**(« 1 avenue de Finlande », 25000 , 1250)

L'entrepôt situé au 1 avenue de Finlande a un volume de 25000 m<sup>3</sup> et une superficie de 1250 m<sup>2</sup>.

**COLIS**(« COL001 », « Paris », « Rennes », 1, 1, 15, 2.5, « 1 avenue de Finlande »)

Le colis dont l'identifiant est COL001 a pour ville d'origine Paris et part de l'entrepôt situé au 1 avenue de Finlande. Sa destination est Rennes. Il est prioritaire, livré à domicile, a coûté 15 euros à l'acheteur et émettra au total 2.5 grammes de CO<sub>2</sub>.

**ITEM**(« Stylo Bic Cristal Bleu », 120, NULL)

Le stylo Bic Cristal Bleu a une masse de 120 grammes et n'est pas taxé.

Note : on aurait aussi pu mettre 0.0 à la place de null.

**Contient**(« COL001 », « Stylo Bic Cristal Bleu », 20)

Le colis dont l'identifiant est COL001 contient 20 stylos Bic Cristal de couleur bleu.

MOYEN\_DE\_TRANSPORT(« Avion », 3.7)

MOYEN\_DE\_TRANSPORT(« Routier », 0.8)

L'avion est un moyen de transport qui coûte 3 euros et 70 centimes au colis pour chaque kilomètre parcouru.

Le transport routier en est un autre qui lui coûte 80 centimes par kilomètre.

TRANSPORTEUR(112233445, « ExcedVitesse », « 15 rue de L'amende »)

La société de transport identifiée par le SIREN 112233445 et nommée ExcedVitesse est basée au 15 rue de L'amende.

Transport(« COL001 », 123456789, « Avion », 458)

Transport(« COL001 », 112233445, « Routier », 14)

Le colis identifié par COL001 a d'abord été transporté par la société dont le SIREN est 123456789 par avion sur 458 kilomètres. Il a ensuite été pris en charge par une autre société dont le SIREN est 112233445 et a parcouru 14 kilomètres via transport routier.

Note : certains colis peuvent apparaître plusieurs fois dans la table, parfois même avec le même SIREN, tout comme certains peuvent n'apparaître qu'une seule fois. La cohérence voudrait qu'un transport se termine systématiquement par voie routière, à moins bien sûr qu'un pilote d'avion vise suffisamment bien pour atteindre un jardin depuis les cieux.

## 6.Requêtes

### a) une sélection avec projection

- Français : tous les moyens de transport qu'a nécessité l'acheminement du colis d'identifiant COL001

- Algèbre :  $\pi_{\text{Type}}(\sigma_{\text{ID\_Colis}=\text{« COL001 »}}(\text{Transport}))$

- SQL : `SELECT Type FROM Transport WHERE ID_Colis = « COL001 »`

- Explication si besoin : on sélectionne les moyens de transport qui ont géré le colis COL001 puis on projette sur leur type.

### b) une jointure

- Français : tous les noms d'items en circulation, leur ville d'origine à chacun et la quantité en transit

- Algèbre :  $\pi_{\text{Denomination, Ville\_Origine, Quantite}}(\text{COLIS} \bowtie \text{Contient})$

- SQL : `SELECT Denomination, Ville_Origine, Quantite FROM COLIS NATURAL JOIN Contient`

### c) une différence

- Français : tous les colis qui n'ont pas encore été transportés et qui sont dans la BDD

- Algèbre : `COLIS – Transport`



- SQL : `SELECT * FROM COLIS WHERE ID_Colis NOT IN (SELECT ID_Colis FROM Transport);`

Note : comme dans MySQL la commande EXCEPT du cours n'existe pas, on est obligé de trouver des alternatives. Ainsi la formule d'algèbre relationnelle  $COLIS - Transport$  est la plus simple, mais celle qu'on effectue vraiment dans SQL est plutôt  $\sigma_{COLIS.ID\_Colis \notin (\pi_{ID\_Colis}(Transport))}(COLIS)$

**d)** une requête avec une condition sur des valeurs nulles

- Français : tous les items qui ne sont pas taxés

- SQL : `SELECT * FROM ITEM WHERE Taxe IS NULL OR Taxe = 0.0;`

Rmq : on doit chercher les valeurs de Taxe qui sont null ou bien égales à 0.0.

**e)** une requête avec une condition contenant ALL et une sous-requête

- Français : le nom des transporteurs dont le SIREN est égal à tous les SIREN dans Transport pour lesquels la distance est égale à 1000.

- SQL : `SELECT Nom FROM TRANSPORTEUR WHERE Numero_SIREN = ALL (SELECT Numero_SIREN FROM Transport WHERE Distance = 1000);`

- Explication : pour que le résultat de la requête soit non-vidé, il faut qu'un transporteur ait un SIREN égal à tous les SIREN dans TRANSPORT qui ont pour Distance la valeur 1000. Cette requête fût très difficile à construire.

**f)** une moyenne sur l'intégralité d'un attribut

- Français : la moyenne de toutes les distances référencées dans la table

- SQL : `SELECT AVG(Distance) FROM Transport;`

**g)** deux requêtes différentes de regroupement avec calcul, dont l'une avec HAVING

1<sup>ère</sup> requête :

- Français : la moyenne de la distance parcourue par les colis venant d'une certaine ville

- SQL : `SELECT Ville_Origine, AVG(Distance) FROM Transport NATURAL JOIN COLIS GROUP BY Ville_Origine;`

2<sup>ème</sup> requête :

- Français : le nombre de colis originaire de chaque ville

- SQL : `SELECT Ville_Origine, COUNT(*) FROM COLIS GROUP BY Ville_Origine HAVING COUNT(*) > 0;`

- Explication : cette requête calcule le nombre de colis qui ont telle ville comme attribut Ville\_Origine. Si deux colis viennent de Toulouse, alors la requête rendra le n-uplet (« Toulouse », 2). Le HAVING permet d'éliminer toutes les villes référencées qui n'ont pas de colis en circulation.

## 7.Conclusion

Tout au long de l'UE et au cours de ce rapport, nous avons appris à construire une base de données, en comprendre les composantes et schématiser son contenu et ses relations. Nous avons aussi appris les bases du SQL, ainsi que toutes les règles qui concernent les bases de données et le peuplement des tables. La contrainte du développement durable nous a quelque peu freinés dans le choix du sujet mais au final nous sommes tous les deux satisfaits du sujet comme du travail rendu. Bien évidemment, les services de gestion de colis utilisent des bases mille fois plus efficaces et optimisées que notre prototype, mais cette base de donnée est tout de même largement améliorable et a une utilité non-négligeable compte tenu de la croissance folle des flux de colis, surtout depuis la pandémie de 2020 en France.

## 8.Annexe

Dans le .zip que nous avons soumis, vous trouvez les fichiers .sql permettant de créer la base de données au complet et de la peupler avec les quelques n-uplets que nous avons utilisés sur phpMyAdmin. Pour finir, nous vous remercions d'avoir lu notre rapport.