

# **ELABORATION D'UNE BASE DE DONNÉES ROUTIÈRES**

## **CONTEXTE**

L'étude s'inscrit dans le cadre de l'opération de recherche du LCPC Métrologie du Trafic et des Trajectoires (MTT) et plus particulièrement dans l'objectif de créer une « Base de données de trafic microscopiques ».

## **CONTENU**

Cette étude donne des pistes de réflexions sur les formats de données et sur l'architecture à mettre en œuvre dans la perspective de partager des données entre fournisseurs et demandeurs.

Nous avons tout d'abord mener une réflexion sur un format d'échange de données à adopter avant de s'intéresser aux données à stocker. Nous proposons ensuite une architecture d'application Web basé sur un SGBDR (système de gestion de base de données relationnelle) où fournisseurs et demandeurs peuvent déposer ou recueillir des données microscopiques de trafic.

## **FORMATS D'ÉCHANGES DE DONNÉES.**

L'étude est plus particulièrement focalisée sur les avantages et inconvénients respectifs de 3 types de format de données couramment utilisés non propriétaire:

- le format CSV ou texte permettant l'échange de données sous forme de tableau,
- le format SQL utilisée par des logiciels manipulant des bases de données relationnelles (logiciels SIG, systèmes de gestion de bases de données (SGBD)),
- le format XML, format d'échange couramment utilisée pour l'échange de données sur Internet et entre systèmes d'information.

## **ELABORATION DE LA BASE**

Les données individuelles sont des données relatives à des véhicules circulant sur une section courante recueillies par différents capteurs implantés dans la chaussée.

Les principaux attributs d'une données individuelle recueillie sur une section courante sont :

- heure de passage du véhicule;
- voie de circulation
- vitesse du véhicule;
- longueur du véhicule;
- attributs relatifs au nombre d'essieux et poids du véhicule;

Ces attributs sont utilisés pour constituer ceux de la base des données.

Attribut	Type de données	Commentaires
----------	-----------------	--------------

Horodate	Date/Heure	'yyyy-mm-dd hh:mm:ss.nnn' ou 'yyyy-mm-dd hh:mm:ss.nnn+hh'
Voie de circulation	Entier	Clé étrangère
Vitesse instantanée	Entier	En km/h
Longueur	Entier	En décimètres

Les attributs qui peuvent être calculés à partir des données de base décrites précédemment ne sont pas stockés dans la base<sup>1</sup>. Il est toutefois possible de développer des fonctions (procédures stockées) exécutées directement par le SGBDR qui calculent ces attributs à la demande de l'utilisateur ou d'un client se connectant à la base de données.

Entre le fichier brut issu de la station de mesure et la mise à disposition de données individuelles depuis la base de données, le contrôle des données peut s'effectuer à différents stades de l'importation :

- lors de la conversion du fichier de données brutes en fichier d'échange de données;
- lors de l'insertion des données dans la base de données;
- après insertion des données dans une base de données.
- 

Les procédures stockées sont des fonctions pré-compilées, stockées sur le serveur et directement exécutées par la base de données. Ces fonctions peuvent être programmées à l'aide de langage de haut-niveau tel que PL/PGSQL pour le SGBDR PostgreSQL, ou en langage de bas niveau tel que le langage C. Par le biais de ces procédures stockées, nous pouvons définir des fonctions de calcul de distances inter véhiculaires, de temps de présence ou encore de catégories de véhicules à partir des données de base stockées dans la base de données.

La base de données comprendra au minimum 2 tables :

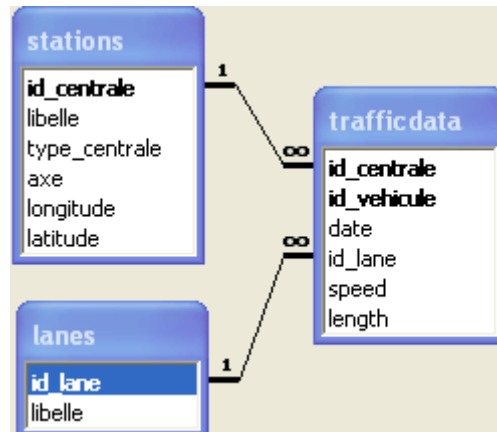
- la table des stations de comptage
- la table de données individuelles

La table des stations de comptage contiendra les caractéristiques de la station de comptage ainsi que celles de la section courante où sont implantées les boucles électromagnétiques.

Concernant les données individuelles, selon le volume de données stockées dans la base, de la puissance du serveur hébergeant le SGBDR et des performances attendues, on peut opter pour une seule table de données ou plusieurs de tailles moindres.

<sup>1</sup> liste non exhaustive d'attributs calculés par le SGBR:

- temps de présence du véhicule;
- distances inter véhiculaires (TIV et DIV);
- classe de silhouette du véhicule;
- contexte de circulation (débit local, données agrégées 6 minutes).



**Relation entre les tables**

L'architecture Web reposera sur le logiciel libre de serveur HTTP APACHE, ce type de serveur étant le plus populaire et le plus utilisé.

Le langage PHP (PHP:Hypertext Preprocessor) sera utilisé comme langage de programmation du côté serveur et PostgreSQL comme système de gestion de bases de données au détriment de MySQL ;PostgreSQL étant plus performant que MySQL dans le traitement et l'exploitation de grosses bases de données.