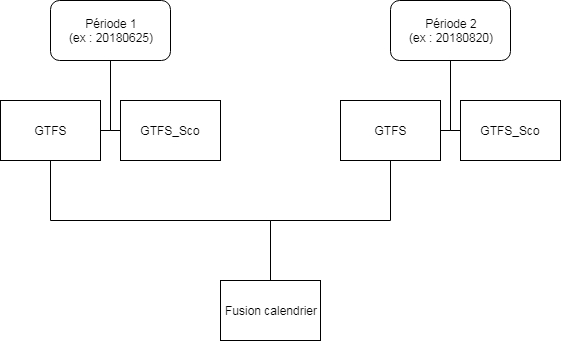
Gestion des GTFS

Le format de fichier GTFS représente le standard dans la diffusion d’information aux usagers. Or, son utilisation pour différents processus crée une multiplication des flux de données produits, chacun avec un ensemble de variations.

À l’heure actuelle, deux types de flux de données sont produits pour chaque assignation, soit un flux contenant l’ensemble des lignes régulières, ainsi qu’un second comprenant en plus les lignes scolaires. Ces fichiers sont spécifiques à chacune des assignations, si bien que les deux fichiers doivent être produits à de multiples reprises dans une année.

Or, lorsque les données de l’assignation suivante sont disponibles, un processus de fusion sur le calendrier est mis en place afin de fournir à l’avance les données sur les différentes plates-formes partenaires (Google, transit, etc.). Cette fusion est faite à l’aide d’un utilitaire produit par Google afin de simplifier différents processus liés aux GTFS. Le résultat de cette fusion est un fichier comportant plusieurs assignations, mais pour lequel certaines entrées existantes dans les deux assignations sont renommées afin d’éviter des conflits. Par exemple, un trip présent à la fois dans l’assignation 1 et l’assignation 2 se verra ajouter un code numérique d’une forme semblable à « merged\_XXXXXX » ou XXXXX est une série de chiffre. Le schéma suivant dresse un portrait de tous les flux existants à un moment où le réseau de l’assignation en cours est en place, et le réseau de l’assignation suivante est déjà connu.



C’est donc 3 GTFS qui peuvent exister pour décrire l’assignation en cours (4 si on fusionne les GTFS scolaires sur le calendrier).

Dans le but d’améliorer l’expérience client dans les différentes interfaces de visualisation, un travail a été fait dans le but de faire une fusion sur le bloc, c’est-à-dire de transformer les lignes en boucle de leur état actuel présentant un aller et retour séparé et de les présenter comme des départs uniques. Le résultat de cette fusion amène donc pour chaque assignation 2 autres GTFS, soit un standard et un scolaire. Or comme chacun des GTFS existant peut être fusion selon le bloc, cette situation amène la possibilité de l’existence de jusqu’à 8 GTFS pour une seule assignation.

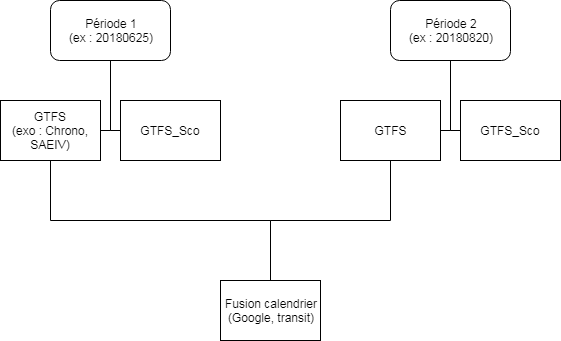
|  |  |
| --- | --- |
| GTFS | GTFS Scolaire |
| GTFS Fusionné sur le bloc | GTFS Scolaire Fusionné sur le bloc |
| GTFS Fusionné sur le calendrier | GTFS Scolaire Fusionné sur le calendrier |
| GTFS Fusionné sur le bloc ET le calendrier | GTFS Scolaire Fusionné sur le bloc ET le calendrier |

La multiplication des flux de données amène une complexité inhérente, alors que les informations fournies par chacun des ensembles de données sont somme toute très semblables. Par exemple, les lignes scolaires étant accessibles à tous les passagers, il y a lieu de se questionner sur la pertinence de ne pas les publier officiellement, auquel cas la complexité pourrait déjà être réduite d’un niveau.

# Problématiques liées

Au-delà de la complexité inhérente à la production et au maintien de multiple flux de données semblables, une autre problématique existe, celle-ci plus importantes. Du fait que de nombreux flux de données existent, il existe un risque réel de désaccord entre différents partenaires qui en font ou pourraient en faire usage.

Un exemple de cette problématique est déjà apparent dans le cadre de la production de données en temps réel par Exo. Lors de la mise en place du SAEIV en début d’assignation, le fichier GTFS de base (ou scolaire) est utilisé. Or, les données fournies aux autres partenaires (google, transit, etc.) changent en cours d’assignation lorsqu’elles sont fusionnées sur le calendrier, changement qui n’est (pour l’instant) pas pris en compte dans le SAEIV. Dans ces circonstances, le résultat net est que les données en temps réel sont produites avec des valeurs ne pouvant parfois plus être réconciliées avec celle fournies à Google dans les cas où certains identifiants sont modifiés avec les suffixes mentionnés précédemment.



Cette situation est appelée à être répétée à chaque assignation lors du processus de fusion sur le calendrier. Aucune piste de solution claire ne peut être soumise à ce moment, puisque l’arrêt de l’envoi des fichiers fusionnés sur le calendrier limiterait la possibilité pour les utilisateurs d’utiliser les calculateurs de trajet au-delà de l’assignation actuelle, alors que le processus interne pour la mise en place des données pour chrono par exo est hors du contrôle de l’équipe de planification développement.