Construire ensemble le nouveau service de décompte ferroviaire

Compte rendu de l'Atelier de Travail de Concertation sur le décompte ferroviaire

28 septembre 2021 – 10h à 12h

Agenda

1. Data Collecting System RTE
    a. Objectifs
    b. Télérelève et orchestration
    c. Contrôles des données
    d. Gestion des doublons
2. API Train Energy Metering
3. Courbes injection et soutirage physiques

Présentation partagée

Retrouvez toutes les informations de la concertation sur le décompte ferroviaire sur le site https://www.concerte.fr/ (GT Données & Comptage)

A l'occasion de cet atelier de travail, SNCF Réseau et RTE ont initié des discussions avec les participants, sur la base de questions qui ont été reprises dans un questionnaire adressé par mail aux participants. Ce questionnaire est disponible sur https://www.concerte.fr/ ou sur demande à RTE-DECOFER-CLIENT@rte-france.com.

RTE et SNCF Réseau rappellent que la présente concertation se déroule en parallèle de la construction du nouveau service, et du développement des nouveaux logiciels associés. Il est nécessaire que les EF et GI participant à la concertation fassent part de leurs retours, avis, questions et besoins le plus tôt possible au fur et à mesure de ce processus de concertation, afin que ces éléments puissent être pris en compte au plus tôt par RTE et SCNF Réseau.

Le présent compte rendu vient en complément des slides présentés lors de l'Atelier de Travail.

Diffusable SNCF RESEAU

Synthèse des échanges :

1. Data Collecting System RTE

a. Objectifs

Principaux éléments présentés

- RTE et SNCF Réseau rappellent les objectifs du Data Collecting System RTE :
  - Pour assurer un service de **télé-relève** et de traitement des **données de comptage ferroviaire** de haute **qualité** et grande **fiabilité**, **RTE** met en œuvre son **expertise** acquise sur le **comptage fixe,** en prenant en compte les spécificités du comptage ferroviaire,
    - **grâce au partenariat avec SNCF Réseau**,
    - et à la **participation active** des **Entreprises Ferroviaires** qui assurent l'ingénierie et la maintenance des compteurs
- RTE rappelle les éléments clés concernant la conception du DCS :
  - **Complétude** : aucune donnée ne doit être perdue
  - **Confidentialité** : la confidentialité des données doit être garantie
  - **Disponibilité** : le DCS doit être disponible plus de 99,5% du temps
  - **Supervision** : le DCS doit permettre de superviser le fonctionnement des compteurs, et de qualifier la donnée de comptage.

b. Télérelève et orchestration

Principaux éléments présentés

- En préparation de l'ouverture du nouveau service de décompte ferroviaire, RTE a développé un Data Collecting System permettant **d'assurer la télérelève des compteurs** :
  - **à la nouvelle norme : EN 50463:2017** : le DCS couvre l'intégralité des 8 services de données prévus dans la norme en mode de communication bord / sol FTP et passerelle
  - **à la norme version 2012 et protocole RFF** : le DCS couvre l'ensemble des commandes prévues dans le protocole en mode FTP (Le mode de communication « SMS » n'est pas pris en charge (similaire à SOCLE))
- Il existe deux mécanismes d'échanges de données entre le **DHS** et le **DCS** :
  - **Mode transfert automatique** : le DHS envoie automatiquement des données au DCS lorsqu'elles sont disponibles avec une périodicité paramétrée dans le DHS (a minima une fois par jour).
  - **Demande / réponse** : le DCS fait une demande au DHS. Le DHS prend en compte la demande, réalise l'action et émet une réponse (en renvoyant les données demandées le cas échéant)
- Le DCS RTE réalise l'orchestration de la télérelève. **L'orchestration est l'intelligence qui décide quand envoyer des demandes au DHS.** Ces demandes viennent **en complément des transferts automatiques**. Cette orchestration est nécessaire pour plusieurs raisons :
  - RTE souhaite obtenir la donnée au plus tôt afin d'accélérer le calcul des décomptes et la mise à disposition des données

Diffusable SNCF RESEAU

- Envoyer une requête toutes les 1h permet de maximiser les chances de coïncider avec un moment où l'EMS est actif (évite l'effet tunnel)
- Les transferts automatiques peuvent dysfonctionner
- Seule une partie des données que peut transmettre le DHS est concernée par les transferts automatiques (BE AUTO en 2012, Automatictransfer en 2017)

- Pour tous les EMS au statut « actif », le DCS réalise le processus en régime établi « orchestration en vie courante » :
  - Toutes les **1h**, le DCS envoie une demande aux compteurs, en spécifiant une période temporelle (date de début et date de fin)
    - Ces demandes concernent, pour le protocole RFF, le « bloc d'énergie », et pour la norme 2017, les services « CEBD, EventSet, ReadingBlock ».
    - *Nota – pour les compteurs norme 2017, toutes les **24h**, le DCS envoie aussi une requête pour **States**, **Asset Data**, et **CommunicationConfig (si les services de données sont gérés par le DHS)***
- **Optimisation de l'orchestration** : les modalités suivantes permettent d'optimiser les ressources et de limiter les doublons :
  - Avant d'envoyer une requête, le DCS **vérifie** systématiquement que les **données qu'il s'apprête à demander n'ont pas déjà été reçues** :
    - Soit suite à une requête (automatique ou manuelle) précédente ;
    - Soit suite à un transfert automatique (à l'initiative de l'EMS).
  - Pour les compteurs à la norme 2017, DECOFER n'envoie les requêtes **que si le service concerné est géré par l'EMS**. DECOFER connait la liste des services gérés par chaque EMS grâce au **processus de synchronisation**, qui a lieu au moment de l'inscription de l'EMS dans DECOFER.
- **Orchestration : extension de la période temporelle de la requête**
  - Toutes les 1h, avant d'envoyer la requête au compteur, le DCS détermine la plage temporelle à demander. Par défaut**, il s'agit de la dernière heure écoulée**.
  - Cette période **peut être étendue** si des **données plus anciennes** sont **manquantes**, par exemple si le compteur n'a pas répondu à une requête précédente.
    - L'objectif est **d'éviter des « trous de télérelève ».**
  - Cependant, il peut arriver que des **données de comptage ne soient pas disponibles** (compteur éteint par exemple), dans ce cas
    - Pour un compteur norme 2017 le compteur répond « *requested data is not available* ».
    - Pour un compteur protocole RFF, on peut le déduire à la réception d'un fichier vide ou dont la période ne contient pas tous les pas de temps demandés.
  - → Dans ce cas, la plage de dates ne sera pas interrogée à nouveau. Le DCS fait la différence entre un *trou de télérelève* et un *trou de données*.

Echanges en séance

- RTE précise que lors qu'un compteur redémarre après avoir été hors tension (locomotive à l'arrêt par exemple), le DCS va interroger l'ensemble de la période pendant laquelle le compteur n'a pas communiqué. Le compteur va alors répondre qu'il n'a pas de données pour la période où il était éteint. Cette période ne sera pas réinterrogée par le DCS : c'est un trou

Diffusable SNCF RESEAU

de données « normal », différent d'un *trou de télérelève*, symptomatique d'un problème de communication bord/sol.

- RTE précise que les compteurs sont généralement prévus pour conserver les données pendant 60 jours.
- RTE rappelle qu'actuellement, SNCF Réseau est responsable du calcul des décomptes via SOCLE, dont les règles de gestion s'appliquent (notamment en ce qui concerne la déclaration des maintenances qui annulent le profilage par type pour les compteurs avec index). Ces règles continueront de s'appliquer pour les compteurs au protocole SAPHYMO. Cependant, à la cible, l'objectif est de ne plus avoir besoin de saisir dans l'outil les périodes de maintenance : le profilage sera appliqué sur la base du journal des circulations transmis par les GI.

c. Contrôles des données

Principaux éléments présentés

- Le DCS réalise un certain nombre de contrôles sur les messages reçus de l'EMS. Ces contrôles permettent de valider la prise en compte des messages et de leurs données dans les traitements de la chaîne, et in fine, dans le décompte des énergies.
- Ces contrôles peuvent amener à invalider certains messages ou certaines données qui ne seront pas intégrées dans le décompte. Ces contrôles sont répartis en trois étapes :

- **Etape 1 : contrôle des fichiers / messages**
  Ces contrôles permettent de vérifier que le fichier / message reçu n'a pas été altéré par un tiers et est bien en provenance du compteur.
  - Si le fichier est chiffré, DECOFER le déchiffre.
  - Si le fichier est compressé, DECOFER le décompresse (en appliquant la méthode décrite dans la norme).
  - DECOFER réalise un contrôle de la signature afin de vérifier la source de la donnée :
    - Pour un compteur au protocole RFF il s'agit du contrôle du CRC
    - Pour un compteur à la norme EN 50463:2017, il s'agit de vérifier la « signature » qui est obligatoire uniquement pour les CEBD
  - DECOFER contrôle le nom du fichier et l'extension du fichier
  - DECOFER contrôle le format du fichier (en le comparant au format défini dans la norme ou le protocole, qui définit la structure).

- **Etape 2 : contrôles de premier niveau du contenu des fichiers / messages**
  - Contrôles liés au compteur :
    - On contrôle l'existence de l'EMS dans la base DECOFER
    - On contrôle le statut de l'EMS dans DECOFER – il doit être « actif »
  - Contrôle de *corrélation* : Pour les messages / fichiers de type réponse reçus de la part du DHS, DECOFER vérifie qu'ils correspondent à une demande préalablement transmises par le DCS
    - Norme 2017 : [header / message ID ; header / correlation ID]
    - Protocole RFF : contrôle du CCD (type de message) et du NR (numéro de requête)

Diffusable SNCF RESEAU

- Contrôle respect de la norme 2017 ou du protocole RFF
  - Compteurs à la Norme 2017 :
    - On contrôle que tous les CEBD ID du fichier sont successifs (ordre chronologique)
    - On contrôle que tous les TimeStamps du fichier appartiennent bien au « measurement interval » spécifié en début de fichier
    - On vérifie que le nombre de CEBD du fichier est inférieur ou égal à la valeur « max CEBD » configurée sur le compteur
    - On vérifie que le nombre de CEBD Blocks du fichier est inférieur ou égal à la valeur « max CEBD Block » configurée sur le compteur
  - Compteurs au Protocole RFF : cas de fichiers en plusieurs tronçons : on vérifie la numérotation des tronçons (de 001 à 999)

- **Etape 3 : identification et contrôle de la valeur des données utiles au décompte**

  - **Utilisation des flags qualité** : D'après la norme EN50463:2017 et le protocole RFF, le DHS doit affecter un « flag qualité » à chaque donnée d'énergie, chaque donnée de géolocalisation et chaque donnée temporelle.
    - Flag qualité donnée d'énergie
      - Les règles de définition de la valeur du flag qualité d'énergie ne sont pas normées et peuvent donner lieu à des traitements différents suivant les constructeurs.
      - DECOFER prévoit, dans une première étape, de **ne pas prendre en compte la valeur de ce flag qualité pour le contrôle des données d'énergie** (identique à SOCLE aujourd'hui).
    - Flag qualité données temporelles
      - DECOFER prévoit, dans une première étape, de **ne pas prendre en compte la valeur de ce flag qualité pour le contrôle des données temporelles**.

      ➔ Suite à l'ouverture du service en 2022 et aux premiers retours d'expérience, RTE prévoit d'étudier et de mettre en place des contrôles de la donnée d'énergie et de la donnée temporelle basés sur ces flags qualité.
    - Flag qualité donnée de géolocalisation
      - DECOFER prévoit de n'utiliser (pour l'affectation GI) **que les données de géolocalisation avec un flag de qualité « ok ».**

  - ***Contrôle des données d'énergie*** *– détection des valeurs aberrantes*
    - DECOFER réalise des contrôles spécifiques sur les données d'énergie.
    - DECOFER affecte un flag « donnée d'énergie aberrante » dans les cas suivants :
      - Energie active consommée en dehors de l'intervalle [0 kWh , 1000 kWh] (sur un pas de 5 min)
      - Energie active régénérée en dehors de l'intervalle [0 kWh , 500 kWh] (sur un pas de 5 min)

Diffusable SNCF RESEAU

- Energie réactive consommée en dehors de l'intervalle [0 kWh , 500 kWh] (sur un pas de 5 min)
- Energie réactive régénérée en dehors de l'intervalle [0 kWh , 100 kWh] (sur un pas de 5 min)

▪ Une fois une énergie qualifiée d'aberrante, elle n'est pas intégrée dans le décompte, et est susceptible d'être profilée :
- profilage par index, si les index sont disponibles et corrects
- Sinon, profilage par circulation dans le cas d'une circulation entière sans données « ok »

○ ***Contrôle des données de géolocalisation*** – *détection des valeurs aberrantes*
▪ Dans certains cas, DECOFER peut affecter un flag « aberrant » aux données de géolocalisation, notamment si
- La longitude et ou la latitude renvoient une valeur aberrante (par exemple 0)
- La localisation renvoyée est située en dehors des polygones des GI

▪ *Pour rappel, la fonctionnalité « affectation GI » a été présentée en atelier de travail le 5/02/2021*

○ ***Contrôle des données temporelles***
▪ On contrôle que chaque TimeStamp est un multiple du TRP (*période de référence, période de temps pour laquelle les données d'énergie sont produites*) qui est de 5 min
- RTE rappelle que DECOFER ne gère que le 5 min (qui est le TRP obligatoire de la norme EN 50463:2017)

Echanges en séance

- RTE précise que le décompte ne sera pas calculé sur un pas inférieur à 5 min, cependant DECOFER va bien intégrer les ReadingBlock au pas 1 minute (données qui seront disponibles en téléchargement via l'IHM DECOFER).
- RTE note bien les remarques exprimées en séance sur les flag qualité de donnée d'énergie : dans un premier temps, il n'est pas prévu d'exclure une donnée du décompte sur la base de ce flag, mais ce sera une évolution possible à étudier dans le futur. Par exemple, un mauvais flag qualité pourrait déclencher un contrôle complémentaire de la donnée via l'index ou l'EventSet. Ces évolutions seront construites en concertation avec les EF.

d. Gestion des doublons

Principaux éléments présentés

- Les équipes RTE / SNCF Réseau ont réalisé des études statistiques sur les données de comptage relevées par SOCLE en 2019. Parmi ces données des doublons ont été détectés.
  - Les résultats montrent des doublons dans les données. La majorité des cas correspond à un même pas de temps télérelevé plusieurs fois (pour un même EMS).
  - Deux éléments vont permettre de gérer ces doublons :

- l'orchestration de la télérelève est conçue pour limiter leur nombre (évite de télérelever plusieurs fois la même donnée)
- Un ensemble de règles de gestion va permettre de traiter les cas de doublons.

- Règles de gestion – **cas de doublons dans 2 fichiers différents**
  Il s'agit du cas le plus fréquent : des pas de temps sont en doublon car ils ont été télérelevés plusieurs fois (on les retrouve dans plusieurs fichiers de télérelève).
  - Cas 1 : *Si les deux enregistrements sont **en tout point identiques**, on conserve le **premier enregistrement.***
  - Cas 2 : *Si les deux enregistrements sont **différents**, on garde celui qui a le **meilleur flag « qualité »***
  - Cas 3 : *Si les deux enregistrements sont différents, et de **qualité identique**, on conserve l'enregistrement le **plus récent**.*

- Règles de gestion – **cas de doublons dans un même fichier**
  Dans des cas plus rares, on retrouve dans un même fichier, deux (ou plus) enregistrements au même pas de temps. Les études ont montré que dans la plupart des cas, il s'agit d'un problème d'horodatage. On retrouve en général un trou de données avant ou après le doublon (ci-dessous cas 1 et 2). La pertinence de cette règle a été vérifiée via une étude statistique sur les boitiers avec index.
  - ***Cas 1 -** Doublon précédant un trou au sein du même fichier. L'EMS horodate mal certains points et les positionne au point d'avant → **Règle :** Modifier l'horodatage pour compléter le trou.*
  - ***Cas 2 -** Doublon suivant un trou au sein du même fichier. Correspond à un problème du DHS qui est arrêté pendant que le compteur continue à compter. A la relance du DHS, un point est créé en plus. → **Règle :** Modifier l'horodatage pour compléter le trou.*
  - ***Cas 3 -** Doublon au sein d'un même fichier, sans trou → **Règle :** une liste de critères est utilisée pour choisir l'enregistrement :*
    - Dé-prioriser les valeurs aberrantes sur l'énergie,
    - Dé-prioriser les valeurs aberrantes sur la géolocalisation,
    - Si différences tensions uniquement, fusionner les enregistrements en conservant les différents niveaux de tension
    - Prendre la source télérelève directe avant les autres sources
    - Prendre la valeur avec la meilleure qualité (flag énergie, flag géolocalisation)
    - Prendre la dernière valeur

2. API Train Energy Metering

Principaux éléments présentés

- L'API proposée constitue un nouveau service de données à la maille des compteurs embarqués. Elle renvoie au client l'ensemble des données de comptage (dans leur état au moment où l'API est sollicitée), correspondant au parc d'EMS de l'EF et à la période temporelle spécifiée.

Diffusable SNCF RESEAU

- Le format d'échange retenu par RTE est le REST/JSON. De plus, afin de préparer les échanges transfrontaliers avec les différents acteurs du marché de l'énergie ferroviaire, la majorité des champs publiés viennent de la norme EN 50 463 (utilisée comme base pour l'IRS 90 930).
- Le recours à cette norme implique le fait qu'en fonction du type de données demandée et de la provenance de la donnée (DECOFER ou SOCLE), il est possible d'avoir des champs non valorisés (ou valorisés avec des null).
- Les données sont disponibles à J+1
- L'API exposée par DECOFER prend en paramètres:
    - Une date de début et une date de fin;
    - Un type de données (brute, ou complétée);
- Les données publiées sont celles associées aux EMS installés sur les engins utilisés par l'entreprise faisant la demande, et dont les dates d'échantillonnage sont comprises entre la période demandée.
- *RTE rappelle en séance les paramètres pris en entrée par l'API*
- RTE met à la disposition des acteurs, sur demande, le guide d'utilisation de cette API
- RTE se tient à disposition des acteurs pour lancer des tests et mettre en place ce flux d'échange de données.

Echanges en séance

- RTE précise que ce sont les EF qui ont accès à cette API (en tant qu'utilisatrices des compteurs), mais que les RE peuvent y avoir accès sur autorisation de transmission de données de l'EF. En ce qui concerne le GI, il a automatiquement accès aux données des compteurs de l'EF lorsqu'elle circule sur son réseau : c'est inscrit dans le nouveau Contrat de Prestation Annexe, et cela ne nécessite pas d'accord séparé.

3. Courbes injection et soutirage physiques

Principaux éléments présentés

- Le calcul actuel tel que réalisé par SOCLE résulte en une courbe d'injection (énergie régénérée ou injectée) la plupart du temps nulle à la maille de l'EF. En effet, SOCLE réalise une étape de **calcul de *l'énergie active nette*** en soustrayant l'injection du soutirage à chaque pas de temps, pour chaque compteur (netting)
    - La courbe de consommation obtenue présente la somme des courbes « nettées » de l'ensemble des compteurs. C'est celle qui est transmise aux RE.
    - La courbe d'injection (énergie régénérée ou injectée) agrégée transmise aux RE est en générale nulle.
- Le calcul que réalisera DECOFER pour la consommation des engins équipés de compteurs résulte en deux courbes (injection et soutirage), et ne prévoit pas de netting :
    - **Soutirage** : somme des énergies actives **consommées mesurées par les EMS**, incluant le **profilage** (par index et par circulation)
    - **Injection** : somme des énergies actives **régénérées mesurées par les EMS**, incluant le **profilage** (par index)

Diffusable SNCF RESEAU

- Pour chaque EF, RTE transmet à son RE une courbe de soutirage et une courbe d'injection
    - *Courbe de soutirage transmise* : somme des énergies actives **soutirées** (**consommées) mesurées par les EMS**, incluant le **profilage** (par index et par circulation), auxquelles on ajoute la **courbe de consommation estimée** (parc **d'engins non équipées**) telle que transmise à RTE par le GI
    - *Courbe d'injection transmise* : somme des énergies actives injectées (**régénérées) mesurées par les EMS**, incluant le **profilage** (par index)
- Ces données d'énergie sont ensuite transformées en puissances moyennes par pas 10 minutes et sont publiées ("données physiques validées")

Echanges en séance :

- RTE précise que les règles de calcul proposées sont cohérentes avec les formules de décomptes de l'ensemble des clients de RTE (hors ferroviaire) et d'homogénéiser les méthodes.
- RTE précise que l'API « Big Metering » utilisée par plusieurs acteurs, renvoie aujourd'hui, pour une EF donnée, les données de comptage pour un point de comptage « fictif » : il s'agit en fait de la consommation de l'EF telle que déclarée à RTE par le GI.

Diffusable SNCF RESEAU