

Acquisition et gestion de données pour le phénotypage animal

Séminaire BigData – 10 mars 2016



SOMMAIRE

- ❖ Définitions
- ❖ Acquisition des données pour le Sicpa
- ❖ Gestion des données du Sicpa
- ❖ Un système d'informations cohérent
- ❖ Exploitation des données

Définitions

❖ Cati Sicpa

- ❖ Cati : Centre Automatisé de Traitement de l'Information
- ❖ Sicpa : Systèmes d'Informations et Calcul pour le Phénotypage Animal
- ❖ Missions : Mettre en place des outils pour intégrer des données de phénotypage animal depuis l'acquisition jusqu'au traitement statistique et à la valorisation de ces données
- ❖ Composé d'informaticiens des départements GA et PHASE

❖ Phénotypage animal

- ❖ Mesure des différents caractères des animaux
 - ❖ Caractéristiques visibles : taille au garrot, réactions de stress, comportement social ou sexuel, ...
 - ❖ Caractéristiques « moins visibles » : qualité de la viande, concentration sérique en hormone, ...
- ❖ Objectif : Caractérisation de sous populations

❖ Phénotypage haut débit

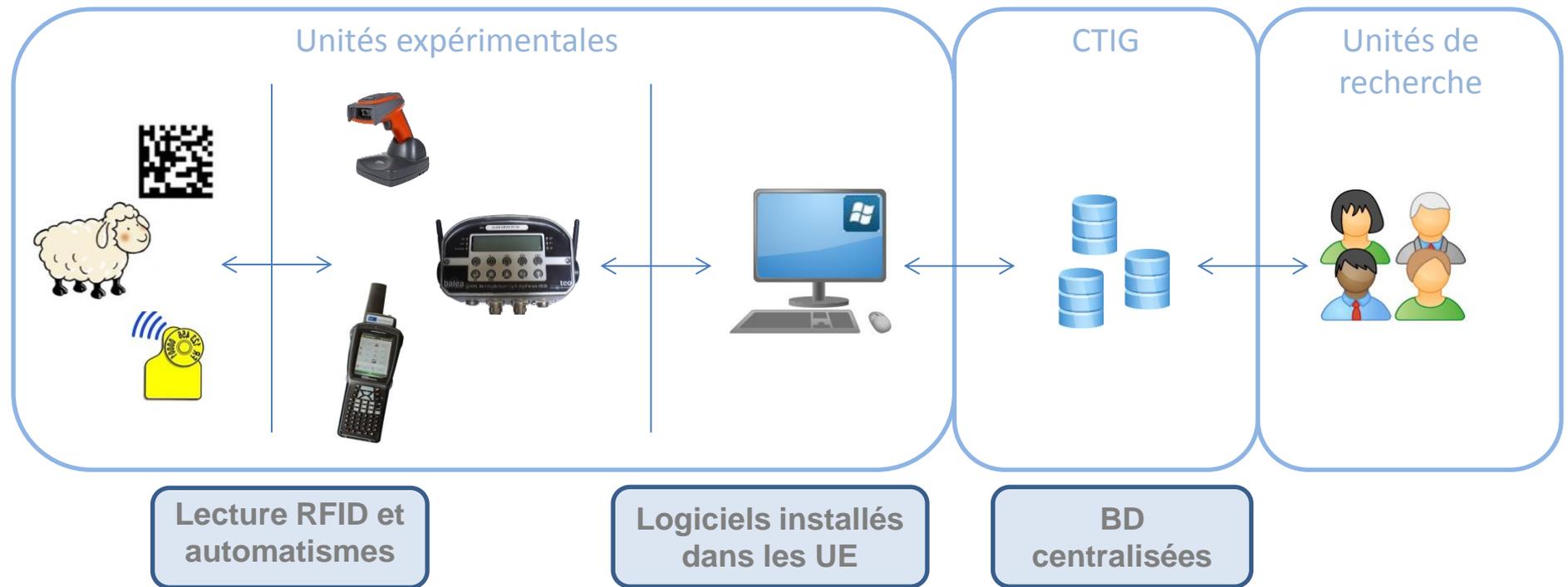
- ❖ Mesures faites de façons standards, répétées et fréquentes
- ❖ Objectif : Vision zootechnique et biologique exhaustive des animaux présents dans les unités expérimentales



Acquisition des données pour le Sicpa

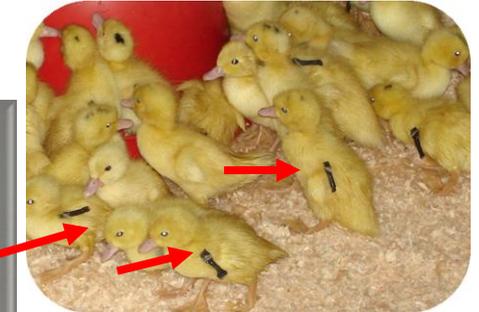
SICPA : Système d'Informations et de Calcul pour le Phénotypage Animal

❖ Chaîne d'acquisition : de la donnée brut à sa valorisation



Acquisition des données pour le Sicpa

❖ Identification des animaux par lecture RFID



- ❖ Technologie norme ISO
- ❖ Numéro d'identification officielle unique assuré par l'EDE
- ❖ Fiabilité de la lecture
- ❖ Produits du commerce ou adaptés au besoin
- ❖ Diminution des risques pour la récupération du numéro

Acquisition des données pour le Sicpa

Utilisation des réseaux sans fil

- ❖ Agnelage BALEA : automate de pesée

Lecteur RFID



Pesée de l'agneau

Pesée de la brebis



Base de données sur serveur local

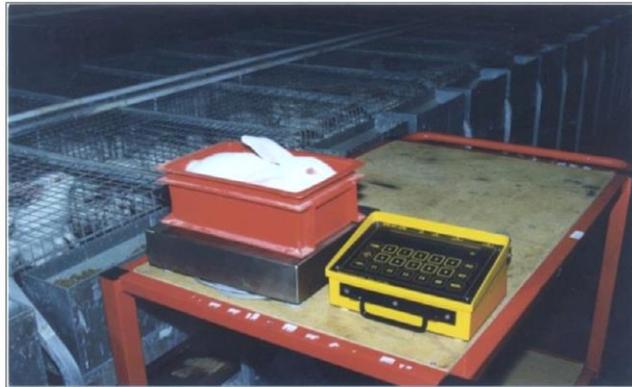


Données enregistrées :

- Date, numéro RFID
- Pesées
- Etat de santé
- Nombre et sexe des agneaux
- Défauts de naissance
-

Acquisition des données pour le Sicpa

❖ Automates de pesée



```
LISTE 1  ESSAI  DATE ?  2  12/04/07
LOT:00 CPESSE:00 REFORME:00 ID:3 CHIP
BATTERIE:11.8v          CMPT:0000

ANIMAL  1070003
ELEVEUR 31102541
SEXE    FEMELLE
POIDS   027,52kg
GMQ     0317g

[ OUI ] [ CPES ] [ SEXE ] [ REF ] [ NON ]
```

Données enregistrées :

- Date, numéro RFID
- Poids
- Sexe
- GMQ (Gain Moyen Quotidien)

Mais aussi :

- Algorithme pour maîtriser les erreurs de mesure

Acquisition des données pour le Sicpa

❖ DAC : Distribution d'Aliment Automatisée / DH2O / DLait



Données enregistrées pour chaque animal:

- Numéro Rfid
- Poids de la nourriture
- Date et heure
- Durée de la visite
-

Etudes :

- Test efficacité de l'aliment
- Qualité nourriture / qualité de la viande
- Comportement alimentaire
- ...

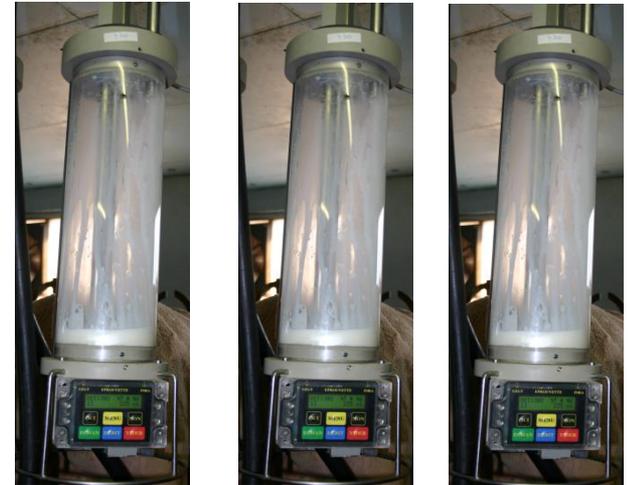
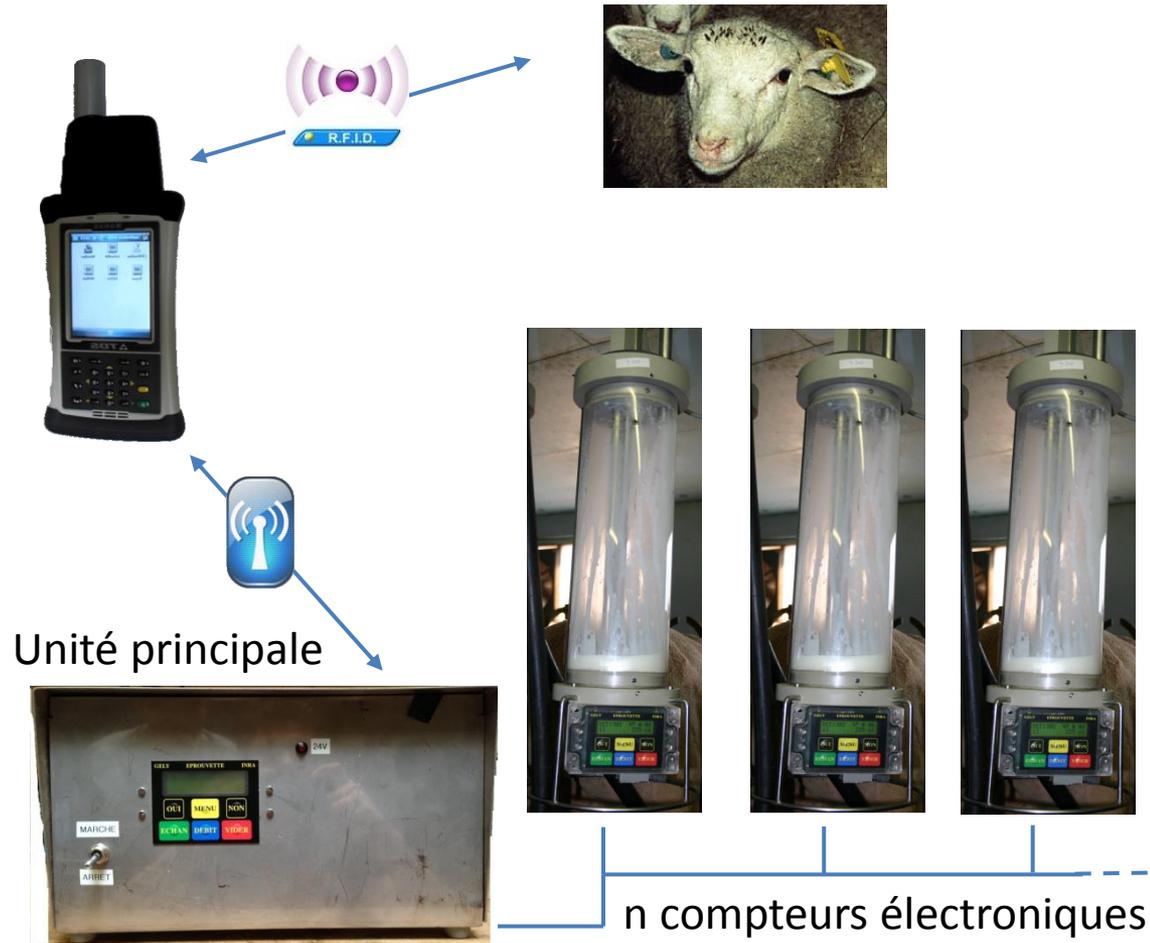
Ex. de volumétrie pour le DAC canards :
pour 20 animaux :
150 000 enregistrements / jour
Fichier à transférer pour BD : 25,5 Mo

Acquisition des données pour le Sicpa

❖ Contrôle laitier

Données enregistrées :

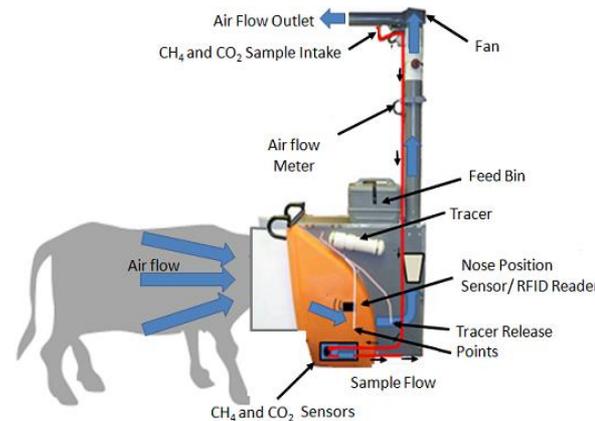
- Date, numéro Rfid
- Quantité de lait
- Temps de latence
- Mesure du débit
- Durée de la traite
-



n compteurs électroniques

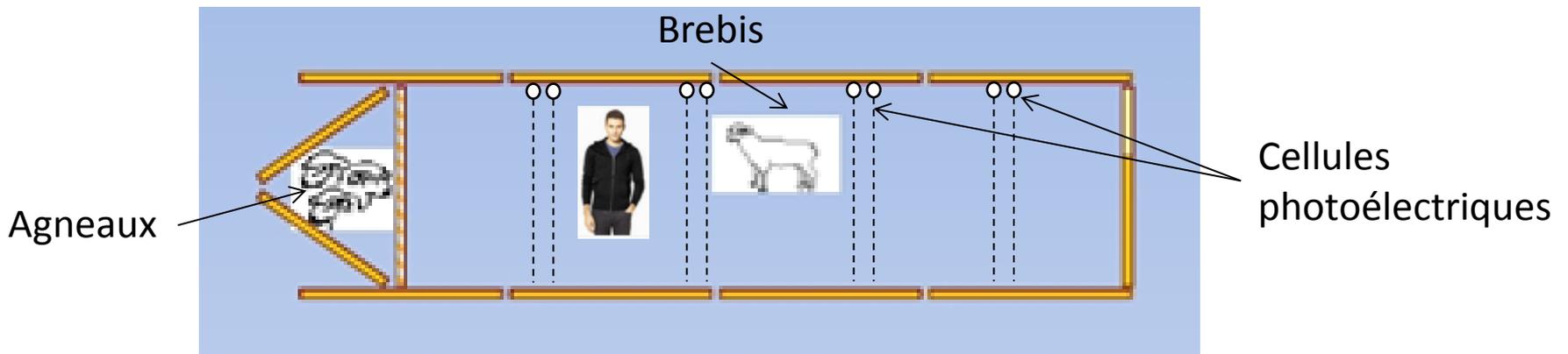
Acquisition des données pour le Sicpa

- ❖ GreenFeed : Mesure des dégagements de méthane et du CO₂ chez les vaches



Système développé par C-Lock Inc.

- ❖ Test de comportement Ovin



Acquisition des données pour le Sicpa

- ❖ Tout cela produit beaucoup de données...
- ❖ Ces données sont enregistrées dans plusieurs bases de données
 - ❖ Centralisées et hébergées au CTIG (Centre de Traitement de l'Information Génétique)
 - ❖ MySQL ou Oracle



Gestion des données pour le Sicpa

❖ Gestion technique des élevages

❖ Données liés à la gestion des élevages

❖ 6 bases espèces

❖ Des données depuis 1970

❖ Quelques chiffres :

❖ 1,6 million d'animaux

❖ 3 millions de pesées (200 000 / an)

❖ 6 millions de données de traites (670 000 / an)



Gestion des données pour le Sicpa

❖ Données expérimentales

❖ Données liées aux protocoles établies par les scientifiques

❖ Base de données : Sicpa Expérimentations

❖ Base de données multi espèces

❖ Création de la base : 2003

❖ Des données depuis : 2005

❖ Taille : 67 Go

❖ Quelques chiffres :

❖ 600 millions de mesures (475 millions en 2015)



Gestion des données pour le Sicpa

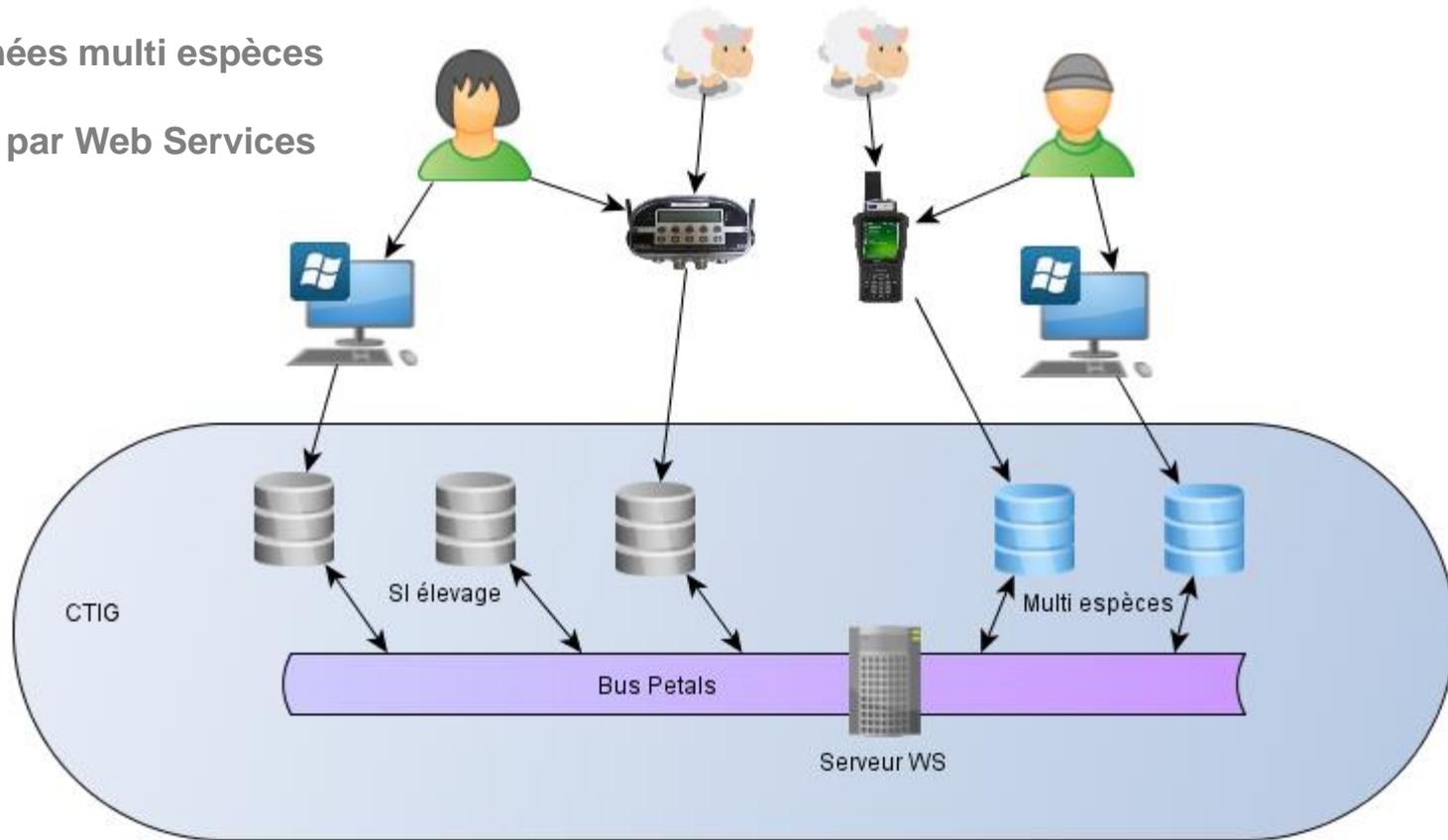
❖ Sicpa Sanitaire

- ❖ Base de données multi espèces
- ❖ Création de la base : 2014
- ❖ Des données depuis 2007
- ❖ Taille : 150 Mo
- ❖ Quelques chiffres :
 - ❖ (48 000 animaux)
 - ❖ 260 000 évènements
 - ❖ 275 000 traitements



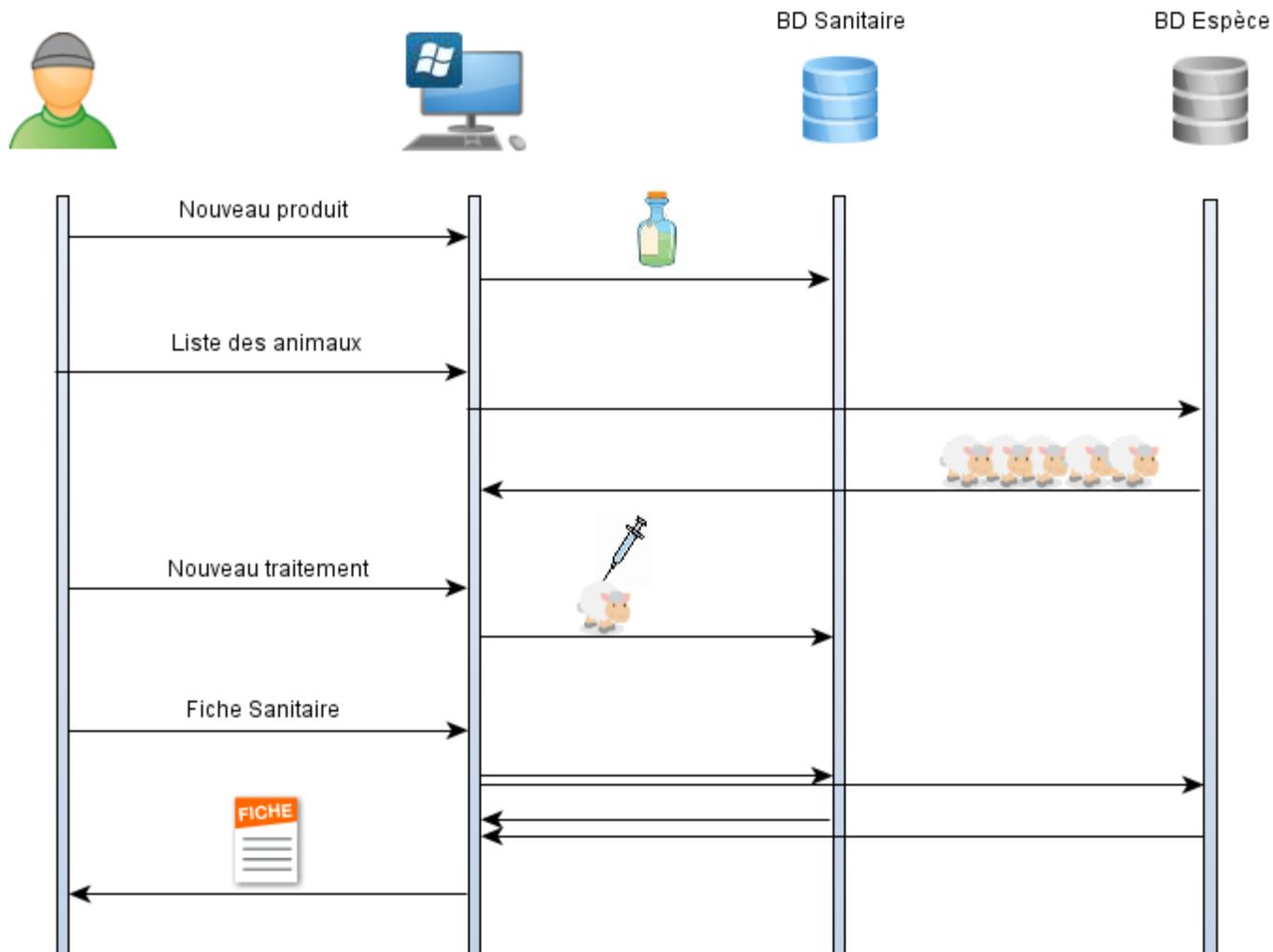
Un système d'information cohérent

- ❖ 6 bases de données « espèce »
 - ❖ SI Poisson en développement
- ❖ 2 bases de données multi espèces
- ❖ Communication par Web Services



Un système d'information cohérent

❖ Exemple avec le Système d'informations Sanitaire



Exploitation des données

❖ Actuellement

❖ Plusieurs solutions existent :

- ❖ Résultats synthétiques depuis les applications
- ❖ Accès direct à la base via un compte de consultation et des vues dédiées
- ❖ Export des données (format csv ou à destination de bases SAS ou R)
- ❖ Interrogation via les Web Services
- ❖ Application web pour la consultation des données

Exploitation des données

- ❖ En cours de réflexion

- ❖ Problématiques :

- ❖ Actuellement les données sont stockées sur des bases de données relationnelles
 - ❖ Fiabilité et robustesse des SGBDR
 - ❖ Expérience des chefs de projet
- ❖ Mais augmentation du volume de données à prévoir très importante
 - ❖ Inquiétude sur les temps de réponse à venir



Exploitation des données

- ❖ En cours de réflexion

- ❖ Plusieurs pistes :

- ❖ Réflexion autour des technologies NoSQL

- ❖ Rapprochement avec le Cati Codex (phénotypage végétal) et son expérience sur MongoDB

- ❖ Etude comparative de temps de réponse

- ❖ Demande de croiser les données sanitaire avec les données élevage

- ❖ Réflexion sur les droits d'accès aux données

- ❖ Lien avec l'Open Science Data

- ❖ Projet de compléter l'ontologie ATOL (Animal Tait Ontology for Livestock)

- ❖ Entrée via les données sanitaires