

Projet Siasp-ONP

Dossier d'architecture



Descriptif du Document

Auteurs : K. Tachfint
Statut du document : en cours
Version du document : 3.0

HISTORIQUE DES VERSIONS SUCCESSIVES

N°	Date	Auteur(s)	Modifications
0.1	17/12/2013	K. Tachfint	Initialisation du document
0.2	06/01/2014		Version en relecture CDAP - CPS
0.3	07/01/2014		1 ^{er} envoi cellule architecture
1.0	19/02/2014	K. Tachfint	Corrections suite à la réorientation du périmètre
1.1	20/02/2014		Version en relecture CDAP - CPS
2.0	24/02/2014		2 ^{eme} envoi cellule architecture
2.1	07/05/2014		Ajout du schéma d'architecture applicative
2.2	13/05/2014	K. Tachfint	Rédaction de la partie 3. l'architecture applicative + mise à jour dossier suite aux nouvelles orientations du projet
2.3	15/05/2014		Version en relecture CDAP - EPS
3.0	21/05/2014		3 ^{eme} envoi cellule architecture

SOMMAIRE

1	Présentation générale du projet.....	4
1.1	1.2. Préambule.....	4
1.2	Programme d'évolution du Siera.....	4
1.3	Une contrainte forte : l'application Siasp actuelle.....	4
1.4	Projet Siasp-ONP.....	7
2	Architecture fonctionnelle.....	9
2.1	Description des Acteurs.....	9
2.2	Fonctions de l'application.....	10
2.2.1	<i>Vue d'ensemble.....</i>	<i>10</i>
2.3	Regroupement des fonctions en sous-systèmes fonctionnels.....	10
2.3.1	<i>Lien avec la matrice générique du processus statistique (GSBPM).....</i>	<i>10</i>
2.3.2	<i>Organisation en domaines et sous-domaines fonctionnels.....</i>	<i>11</i>
2.4	Interactions avec les autres systèmes.....	12
2.4.1	<i>Données utilisées par Siasp-ONP.....</i>	<i>12</i>
2.4.2	<i>Échanges avec les opérations de libre-service.....</i>	<i>13</i>
2.4.3	<i>Données fournies par Siasp-ONP.....</i>	<i>14</i>
2.5	Diagramme de flux.....	14
2.5.1	<i>Synthèse des flux entrants et sortants du système.....</i>	<i>15</i>
2.6	Exigences et contraintes technico-fonctionnelles.....	16
2.6.1	<i>Exigences (version provisoire sur la base de l'existant).....</i>	<i>16</i>
2.6.2	<i>Les contraintes.....</i>	<i>17</i>
3	Architecture applicative.....	19
3.1	L'architecture applicative générale.....	19
3.2	Accès aux données.....	20
3.2.1	<i>Base de données de production.....</i>	<i>20</i>
3.2.2	<i>Mise à disposition de fichiers.....</i>	<i>20</i>
3.3	L'architecture applicative du module d'accueil-réception.....	21
3.4	L'organisation de l'application Java.....	21
3.5	Les Framework utilisés dans le projet Java.....	22
3.5.1	<i>Utilisation du Framework Spring.....</i>	<i>22</i>
3.5.2	<i>Utilisation du Framework Spring-batch.....</i>	<i>22</i>
3.5.3	<i>Utilisation de l'ETL Talend Open Studio.....</i>	<i>22</i>
3.5.4	<i>3.4.3 Autres outils.....</i>	<i>22</i>
	Annexe 1 : Quels besoins d'accès aux données dans les applications du Siera.....	23

1 Présentation générale du projet

1.1 1.2. Préambule

Suite aux conclusions du rapport Marzin¹, le ministère du budget a acté la simplification du projet ONP. Cette simplification passe par l'arrêt du SI-Paye de l'ONP tout en poursuivant les efforts de modernisation des SIRH ministériels et la mise à niveau technique à iso-fonctionnalités de la chaîne Pay de la DGFIP jusqu'en 2017, chaîne qui continuera à assurer la liquidation de la paye de tous les agents de l'État. Ainsi il n'est plus envisagé d'avoir dans les trois prochaines années un nouveau format de données pour les fichiers de paye des agents de l'État. En conséquence, le projet Siasp-ONP est suspendu.

Faute d'éléments suffisants sur les volumétries et le contenu des fichiers provenant de l'ONP, la partie technique du présent dossier d'architecture est recentrée sur les éléments transversaux du Siera², porté initialement par le projet Siasp-ONP: le module d'accueil-réception et le besoin d'accès aux données de production.

1.2 Programme d'évolution du Siera

Le Siera est un système d'information sur l'emploi et les revenus d'activité. Il a été développé en plusieurs temps³ pour répondre aux missions du département de l'emploi et des revenus d'activité (Dera), à savoir **produire et analyser les statistiques sur l'emploi et les revenus d'activité**, à partir de sources administratives. Il repose actuellement sur huit applications en interactions internes et externes au Siera : DADS 2, Épure 2, Non-Salariés, Particulier-employeur, Estel, Siasp, Ecmoss et Frontal N4DS⁴.

L'activité peut être déclinée en deux macro-processus de production:

- **Produire une information conjoncturelle sur l'emploi salarié et les revenus d'activité ;**
- **Produire une information structurelle sur l'emploi total et les revenus d'activité.**

Un programme d'évolution du Siera se met en place en 2013. Le projet Siasp-ONP fait partie de ce programme.

1.3 Une contrainte forte : l'application Siasp actuelle

Le système d'information des agents du service public (Siasp) existant repose sur la mobilisation de sources administratives. Il s'intéresse à la fois à l'emploi et aux rémunérations, dans une dimension nationale mais aussi territoriale détaillée. Il vise à assurer la meilleure comparabilité possible et une cohérence maximale en matière de revenus et de condition d'emploi des salariés dans les 3 versants de la fonction publique.

La colonne vertébrale de ce système est assurée par l'application Siasp, qui se base sur deux schémas informatiques :

- Un schéma conjoncturel, qui prend en compte les fichiers de paye mensuels des agents de l'État, de leur réception à leur annualisation, en passant par leur traitement trimestriel.
- Un schéma annuel qui prend en compte l'ensemble des sources : fichiers de paye des agents de l'État annualisés, DADS, Défense et TGE (trésorerie générale des agents à l'étranger).

Compte tenu de son ampleur, le projet Siasp avait été scindé en deux lots qui ont été développés par des équipes différentes (lot 1 : structurel au CNIP, lot 2 : conjoncturel et annualisation des fichiers mensuels au CNIO). Les technologies retenues ont été Java 5 et le langage PL-SQL pour les étapes d'agrégation de périodes (Oracle 10).

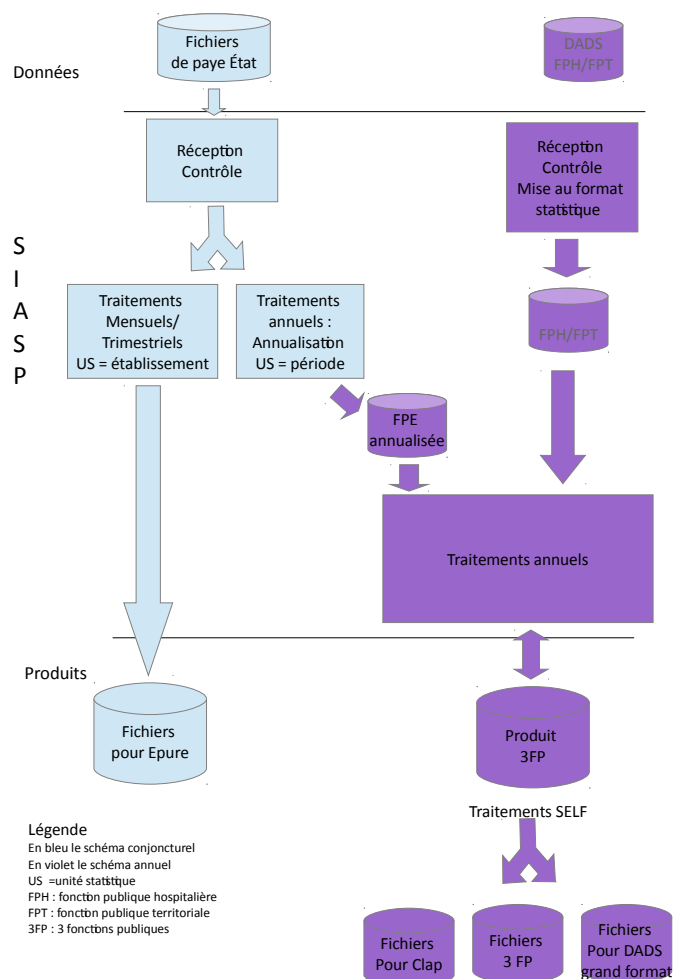
¹ http://www.modernisation.gouv.fr/sites/default/files/fichiers-attaches/rapport_programme_sirh_si-payé.pdf

² Cf. note N° 137/F201 du 30 août 2001 : « Évolution du système statistique sur l'emploi et les revenus d'activité ».

³ Le Siera a été lancé en 2001-2002 et les projets associés se sont déroulés entre 2004 et 2010.

⁴ Cf. note N° 80/DG75-C510/ du 17 mars 2011 : « Étude d'urbanisation sur le domaine du Siera ».

Le projet s'est terminé en 2010, dans un contexte d'évolution de la périodicité de la livraison et du contenu des données sur les agents de la fonction publique d'État (annuelle devenue mensuelle). La cohérence et la complétude du périmètre n'ont pu être assurées durant le projet laissant le développement de certaines fonctionnalités à la charge de l'équipe de maintenance⁵. Ainsi, dans l'état actuel de l'application, il reste encore de nombreuses procédures semi-automatisées et des travaux en self.



Le projet Siasp-ONP consiste en la refonte de la partie conjoncturelle du système d'information existant suite à la mise en place de l'ONP, dont les données vont se substituer aux actuels fichiers de paye de l'État de la DGFIP. **Il faudra donc veiller à la bonne insertion de ce nouveau flux de données dans un existant**, notamment le schéma conjoncturel qu'il faudra alimenter. Sont impliqués dans le processus la section Siasp de la division Exploitation des fichiers administratifs, le pôle Fonction publique à la direction régionale de Bretagne, le service développement du CNI d'Orléans qui assure la maintenance de l'application Siasp et le service production du même CNI pour la production informatique

La forme générale de l'application actuelle est décrite selon un schéma type distinguant 4 blocs d'activités mis en œuvre de façon successive lors du processus de traitement :

Réception-contrôle, Mise au format statistique, Contrôle-appariement, Élaboration des produits

Réception-contrôle des données

Les données proviennent des fichiers de paye fournis par la DGFIP.

La DGFIP fournit un fichier « images de feuilles de paye » et des fichiers de nomenclatures associés. Chaque trésorerie générale fournit un fichier, via les trois centres informatiques de la DGFIP. Les fichiers fournis sont mensuels, ce qui constitue une difficulté pour le processus de production qui avait été initialement conçu pour recevoir des fichiers annuels. Ces fichiers sont communiqués à l'Insee par le service d'échange de fichiers (ORIADE).

Un moteur de réception contrôle est utilisé sur les données DGFIP. Chacune de ces sources subit des

⁵ Cf. Documents de travail N°C2010/07 et C2011/02.

contrôles qui lui sont propres, liés aux informations et aux formats. Ce moteur de réception contrôle est également utilisé par l'application Épure.

Cette partie réception-contrôle des données est réalisée en utilisant des programmes développés en Java 5.

Mise au format statistique

Cette étape comporte plusieurs fonctions : contrôle de l'employeur, du salarié, reconstitution de variables à partir d'autres variables, affectation d'une ligne budgétaire. Ce bloc d'activités est réalisé par des programmes PL-SQL.

Contrôle et appariement

Des contrôles sur le périmètre attendu par rapport au socle Épure sont effectués.

La reprise des codifications douteuses et des incohérences est effectuée par le pôle de Rennes, via une IHM développée en Java 6 et utilisant le Framework Struts 1.

D'autres contrôles sont réalisés en self (SAS) sur la cohérence des données et la réagrégation des périodes.

Enfin, les données sont enrichies par des informations Établissement (SIRET et localisation) en provenance de Clap.

Élaboration des produits

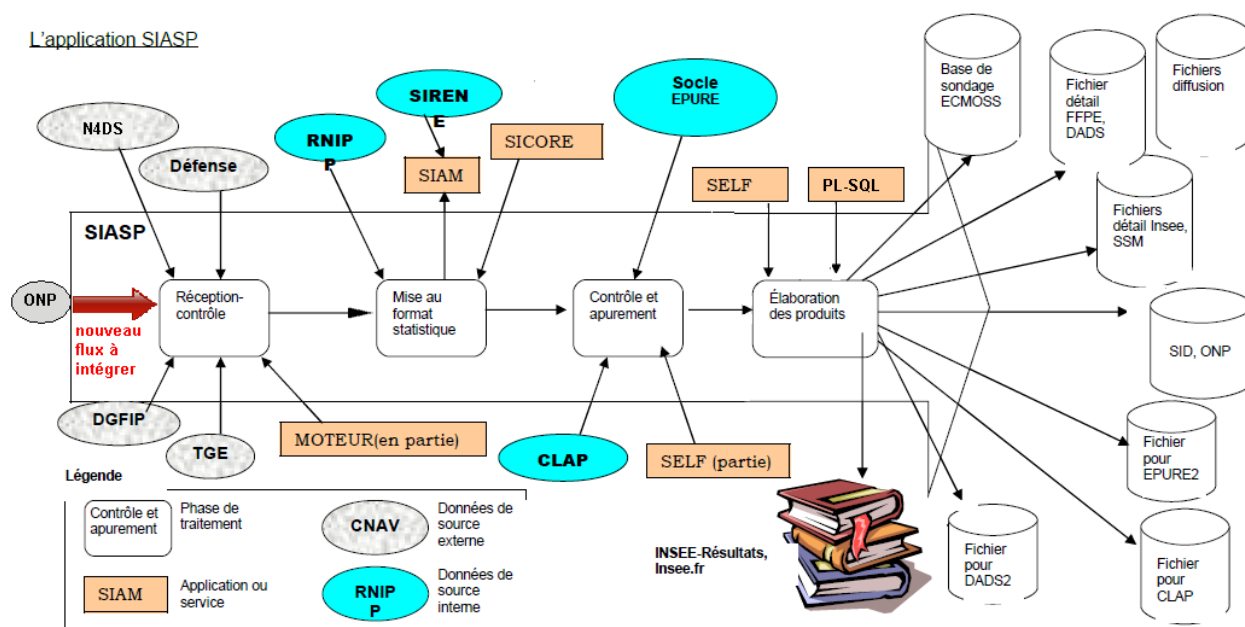
Cette étape est réalisée avec des programmes développés en PL-SQL et/ou en self (SAS).

A partir du schéma conjoncturel Siasp, des produits sont diffusés vers les SSM, l'ONP, l'application Clap(données État), l'application Épure 2 et le schéma annuel Siasp (annualisation des données). Siasp sert également de base de sondage pour constituer l'échantillon d'ECMOSS et celui de l'enquête DGAFP sur les salaires.

A cela s'ajoute un bloc d'activités transversal d'administration de l'application : gestion des accès, suivi de production, journalisation des événements.

Le schéma d'ensemble, ci-dessous, est une représentation des grandes fonctions métier de l'application, donnant en outre le périmètre entre les fonctionnalités réalisées en libre-service et celles réalisées dans l'application. Le nouveau flux à intégrer en provenance de l'ONP est également représenté. Y sont également décrites les applications clientes.

L'application SIASP



1.4 Projet Siasp-ONP

Une des sources alimentant actuellement Siasp, les fichiers de paye des agents de l'État, va être progressivement remplacée par une autre source : les données de l'Opérateur National de Paye (ONP). Le projet Siasp-ONP consiste donc en l'introduction du flux transmis par l'ONP dans Siasp. Ce projet est motivé par la nécessité de garantir la continuité de production des statistiques trimestrielles et annuelles sur l'emploi de la fonction publique d'État en intégrant l'arrivée progressive de ces données de l'ONP sans interruption :

- a. En assurant la cohérence des données (la transition est prévue sur au moins 5 ans) ;
- b. En respectant les délais et les engagements de diffusion des statistiques prises par le département de l'emploi et des revenus d'activité.

Dans ce cadre, le contenu actuel des produits, trimestriel et annuel, de l'application Siasp ne sont pas remis en question dans le cadre de ce projet.

Il est déjà acquis que la partie accueil des données de l'application Siasp pour la fonction publique d'État devra évoluer, dans sa dimension trimestrielle et annuelle (i.e. l'annualisation des données mensuelles).

Contexte et contraintes

Le basculement des ministères vers le SI-Paye ONP étant progressif, il est nécessaire de prévoir une alimentation de l'application SIASP simultanément par les deux types de données : fichiers de paye des agents de l'État, transmis par les trésoreries générales, et fichiers de paye ONP. Cette double alimentation devrait durer plusieurs années. Une des difficultés réside dans la possibilité d'assister à la bascule de la paye d'un ministère entier ou uniquement d'une partie de celui-ci et de veiller à ne pas avoir de double compte pendant ces bascules.

Le projet devrait permettre de dégager des solutions réutilisables pour l'accueil des données administratives alimentant le Siera et en particulier les données de la déclaration sociale nominative (DSN) qui devrait remplacer la DADS à compter de 2016.

Le projet devra veiller à assurer le maintien de l'application Siasp, notamment ne pas perturber l'application actuelle qui permet l'alimentation à partir des fichiers de paye transmis par les trésoreries générales (hors périmètre du projet). Ne fait pas partie du projet Siasp-ONP, l'intégration de fichiers autres que ceux transmis par l'ONP. De plus, l'automatisation des opérations réalisées actuellement en Self par la section Siasp ne fait pas partie du projet.

Les options retenues devront permettre de continuer à alimenter les applications clientes de Siasp mentionnées précédemment, sans les impacter.

Objectifs

Les trois objectifs stratégiques sont les suivants :

- ❶ Continuité de la production des statistiques annuelles sur le versant de la fonction publique d'État
- ❷ Continuité de la production des statistiques conjoncturelles sur le versant de la fonction publique d'État;
- ❸ Dissociation des données administratives qui alimentent le Siera et des données statistiques

Les objectifs opérationnels du projet SIASP-ONP sont :

- Prendre en compte ce nouveau mode d'alimentation (flux ONP) dans le système d'information actuel sur les agents des services publics (Siasp);
- Se brancher sur le poste gestionnaire existant pour assurer les opérations de gestion de campagne ONP et toute opération nécessitant une intervention des équipes statistiques ;
- Continuer à livrer en aval les mêmes produits qu'actuellement dans leur forme et leur contenu afin de ne pas avoir à modifier les différentes applications clientes de Siasp ;
- Inscrire ces travaux dans la perspective plus large du programme d'évolution du Siera. Ainsi, les solutions qui seront conçues et développées pour Siasp-ONP, notamment sur la partie accueil-réception des fichiers source, doivent être génériques et pouvoir être réutilisées ultérieurement dans d'autres projets.



Il conviendra de gérer la bascule progressive des ministères à l'ONP en alimentant Siasp avec les deux flux d'information sur la paye des agents de l'État, pendant toute la période de déploiement de l'ONP, en évitant les doubles comptes et en assurant la permanence de l'alimentation de Siasp pour tous les agents de chaque ministère.

2 Architecture fonctionnelle

L'architecture fonctionnelle positionne les acteurs de l'application, les principales fonctions, les regroupements par sous-système ainsi que les contraintes et exigences associées au projet. Ce niveau d'architecture n'est pas particulier à une technologie. Il définit ce qui doit être fait, et non comment cela doit être fait.

2.1 Description des Acteurs

Un acteur, au sens UML, représente le rôle d'une entité externe (utilisateur humain ou pas) interagissant avec le système. Un utilisateur peut être amené à jouer plusieurs rôles vis-à-vis du système et à ce titre être modélisé par plusieurs acteurs.

Les acteurs du système cible sont décrits ci-dessous.

- **Gestionnaire** : il est utilisateur de l'application pour les tâches de gestion courante : suivi de campagne, identification des établissements, ...

Du point de vue organisationnel, les gestionnaires se situent au sein du pôle fonction publique de Rennes à la Direction Régionale de Bretagne. Ils sont actuellement au nombre de deux. Il est prévu qu'à moyen terme, cette activité soit reprise par le centre statistique de Metz.

- **Administrateur d'application** : il est utilisateur direct de l'application à la fois pour les tâches de gestion, qui ne sont pas son cœur de métier mais auxquelles il peut participer, mais aussi pour les tâches d'administration : droits d'accès à l'application, ...

Du point de vue organisationnel, il s'agit du chef de la section Siasp de la division exploitation des fichiers administratifs au sein du Département Emploi et Revenus d'activité.

- **Batch_Siasp** : il s'agit d'un acteur fictif déclenchant l'exécution des traitements ordonnancés (traitements batch), comme les opérations d'ouverture de campagne, de constitution des produits pour les applications clientes.

Du point de vue organisationnel, on peut le ramener au service de production, chargé d'ordonnancer les traitements, et plus précisément, du RIAP du futur site de production de l'application.

Échanges avec les applications :

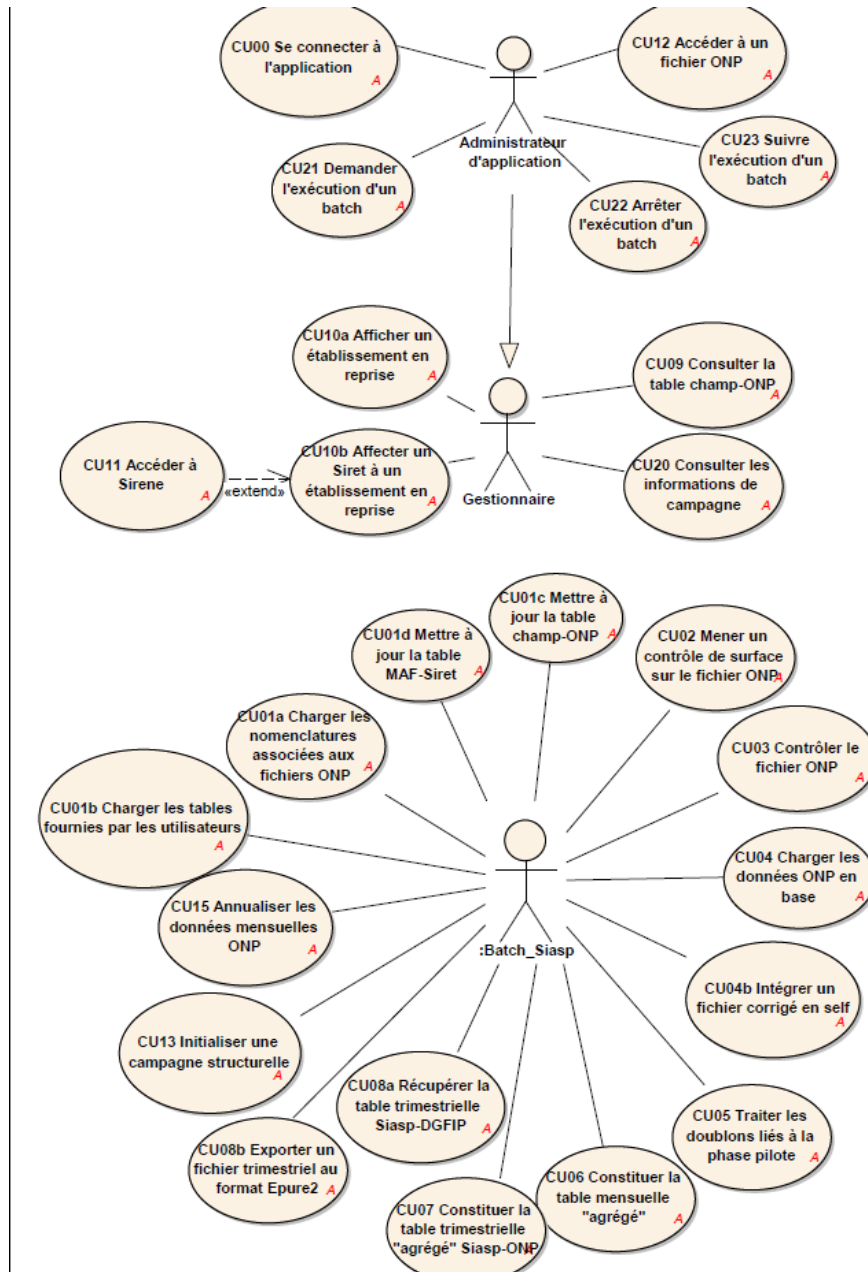
- **Épure 2** : Extension du projet Urssaf aux revenus et à l'emploi
- **Siasp : Système d'information sur les agents des services publics de l'État (partie structurelle)**, également appelé schéma millésimé.
- **Sirene**
- **Le Service d'Échange de Fichiers (SEF)** pour la gestion des transferts de fichiers.

L'application Siasp-ONP a peu d'utilisateurs directs : pas plus de six agents répartis entre le pôle fonction publique à Rennes et la section Siasp à Paris. Les profils de ces utilisateurs sont constitués par des combinaisons des « rôles » décrits ci-avant.

En revanche, l'application a des partenaires internes à l'Insee avec lesquels elle échange de l'information (cf. diagramme des flux).

2.2 Fonctions de l'application

2.2.1 Vue d'ensemble



2.3 Regroupement des fonctions en sous-systèmes fonctionnels

Cette partie permet d'identifier les sous-systèmes fonctionnels de l'application ainsi que leurs interactions avec d'autres applications de l'Insee ou en dehors de l'Institut. Un sous-système peut être perçu comme un système autonome à l'intérieur du projet. Une application peut ne comporter qu'un seul système.

2.3.1 Lien avec la matrice générique du processus statistique (GSBPM)

L'application Siasp-ONP permettra de réaliser deux macro-processus métiers :

- **Produire une information trimestrielle** sur l'emploi salarié dans la fonction publique d'État
- **Produire une information annuelle** sur l'emploi salarié dans la fonction publique d'État

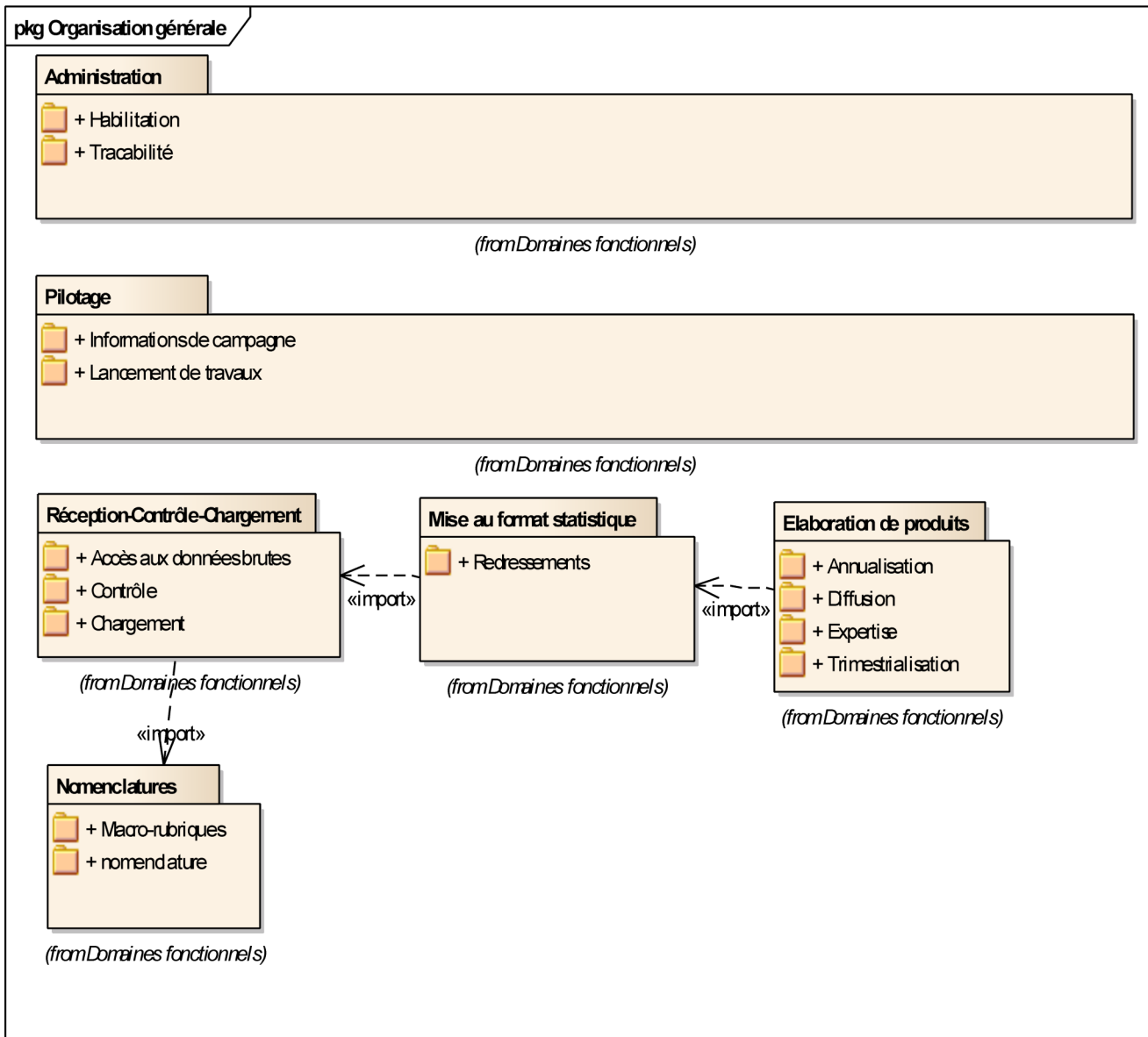
Le découpage de l'application en domaines et sous-domaines fonctionnels s'appuie sur le modèle générique du processus statistique GSBPM. Les éléments répertoriés ci-dessous font donc partie de l'application.

PHASES GSBPM	Sous-processus GSBPM	Fonctions					
4. COLLECTER	Procéder à la collecte	Accueil fichiers					
	Finaliser la collecte	Recopie, renommage, formatage des fichiers					
5. TRAITER LES DONNÉES	Intégrer les données	Transfert fichiers sur serveur spécifique	Chargement des données dans l'application	Repérage et identification des doubles comptes et mise hors champ			
	Classer et coder les données	Codification automatique					
	Contrôler, qualifier, redresser les données	Contrôle de format	Contrôle de présence des informations nécessaires	Contrôle de conformité aux filtres	Contrôle de validité des NIR	Contrôle de validité des SIREN	Comparaison à un socle de référence
	Imputer les données	Redressements manuels					
	Dériver de nouvelles variables ou unités stat	Création de variables à partir d'autres variables					
	croiser avec données DGFIP	Calculer des agrégats	Agrégation de périodes				
6. ANALYSER	Accéder aux données brutes	Préparer la version préliminaire des résultats	Attribuer des indicateurs de qualité				
	Valider les résultats	Contrôle des incohérences internes	Confrontation des valeurs avec d'autres sources				
7. DIFFUSER	Mettre à jour les systèmes de diffusion	Constitution fichier global de diffusion					
	Produire les produits de diffusion	Livraison de fichiers à d'autres systèmes					

2.3.2 Organisation en domaines et sous-domaines fonctionnels

L'application Siasp-ONP se décompose en quatre grands domaines (Réception-contrôle-chargement des données - Mise au format statistique - Contrôle et appariement - Élaboration de produits), eux-mêmes étroitement liés au découpage établi à partir du GSBPM et évoqué au paragraphe 2.2.

Du point de vue de l'architecture informatique, les domaines et sous-domaines fonctionnels organisant l'application sont inspirés de cette représentation.



2.4 Interactions avec les autres systèmes

2.4.1 Données utilisées par Siasp-ONP

Les services et données d'autres applications utilisées par Siasp-ONP sont les suivants :

L'application Siasp existante (sur le champ DGFIP)

1. Siasp-ONP s'appuiera sur les données calculées issues de Siasp existant pour fournir à l'application Épure un fichier trimestriel exhaustif sur le champ de la fonction publique d'État. Siasp-ONP agrégera les données sur le champ DGFIP à celles qu'il aura calculées sur le champ ONP.
2. Siasp-ONP requêtera tous les mois les données chargées dans Siasp existant qu'il croisera avec celles fournies par l'ONP pour s'assurer de la couverture exhaustive du champ de la fonction publique d'État.

Sirene

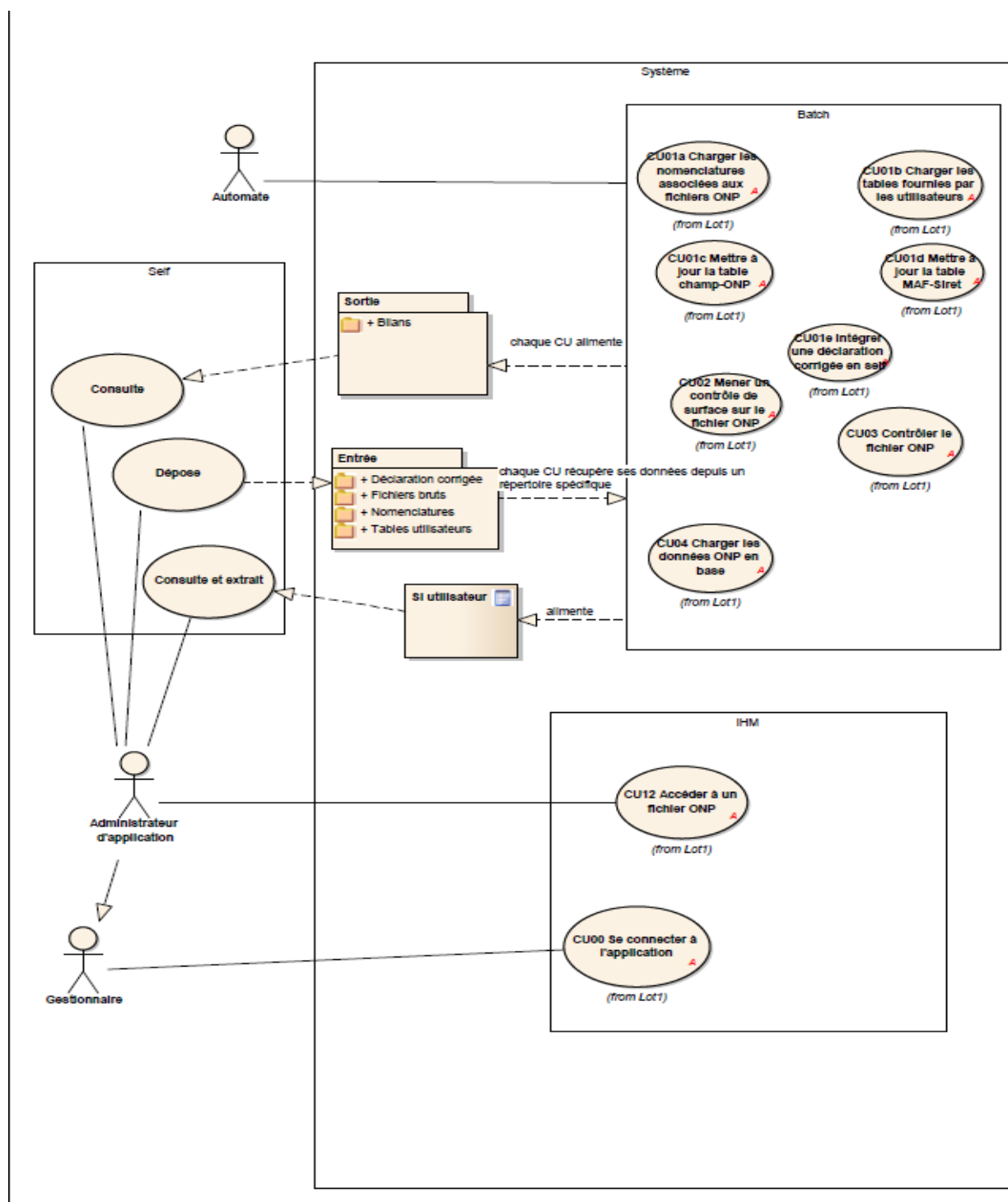
Siasp-ONP utilisera un accès à l'application Sirene utilisée en consultation pour les personnes ne travaillant pas à la gestion Sirene, permettant ainsi des recherches variées et souples sur les établissements.

2.4.2 Échanges avec les opérations de libre-service

L'expérience acquise par le DERA sur le traitement des données administratives montrent que certains traitements ne peuvent être totalement automatisés. C'est pourquoi dans la solution proposée, certains traitements restent du ressort de l'administrateur d'application. Il utilise pour cela les données produites par l'application dans la base de production, réalise des travaux en libre-service, puis fournit à l'application les fichiers résultant de ces travaux afin qu'il soient intégrés dans l'application en production.

Il s'agit donc d'articuler les travaux de l'application et son environnement de production avec l'espace de travail des utilisateurs,

Le diagramme ci-après représente, du point de vue fonctionnel, la façon dont vont se dérouler les échanges entre les traitements réalisés par l'application et les traitements en libre-service. Une note plus détaillée du besoin d'accès aux données/fichiers est jointe en annexe de ce dossier d'architecture.



2.4.3 Données fournies par Siasp-ONP

Les services et données fournis par Siasp-ONP à d'autres applications sont les suivants :

L'application Épure

L'application Épure (Extension du projet Urssaf aux revenus et à l'emploi) est cliente de notre application qui doit lui fournir tous les trimestres un fichier contenant des données économiques (effectifs cumulés, masse salariale,...) par établissement sur le champ public.

Siasp_millesime

Siasp_millesime est le Système d'information sur les agents des services publics (partie structurelle). Il sera alimenté par l'application, Siasp-ONP tous les ans en données socio-économiques structurelles sur le champ des agents de la fonction publique d'État.

2.5 Diagramme de flux

L'application aura de nombreux échanges avec des acteurs externes (autres applications Insee, acteurs hors Insee...).

Au début de la mise en place de l'équipe projet, le calendrier de l'ONP apparaissait comme étant la contrainte majeure du projet, il s'agissait d'être en capacité d'intégrer rapidement les premières livraisons de l'ONP afin d'assurer la continuité de la production des statistiques conjoncturelles.

En conséquence l'équipe projet avait envisagé la constitution d'un prototype (éventuellement jetable) avant l'été 2014, qui permettait d'accueillir, stocker et lire les fichiers transmis par l'ONP. L'idée a par la suite été abandonnée, étant donné le desserrement de cette contrainte de calendrier. La décision actuelle serait d'adopter une stratégie en deux lots:

- Un premier lot dont l'objet serait de traiter seulement la partie « intégration » des données de l'ONP dans Siasp. Sa finalité serait de livrer à Siasp les nouvelles données nécessaires à la poursuite d'un suivi exhaustif. Ce lot permettrait également de traiter les doublons / omissions liées à la double alimentation ONP-fichiers de paie de la DGFIP dans les phases de pilote. Dans ce lot, serait développé un module d'accueil réception des données administratives du Siera.
- Un second lot qui vise à l'amélioration du processus existant notamment sur les phases en aval de la mise au format statistique. L'objectif est de dissocier les concepts statistiques mis en œuvre des données administratives alimentant le système d'information

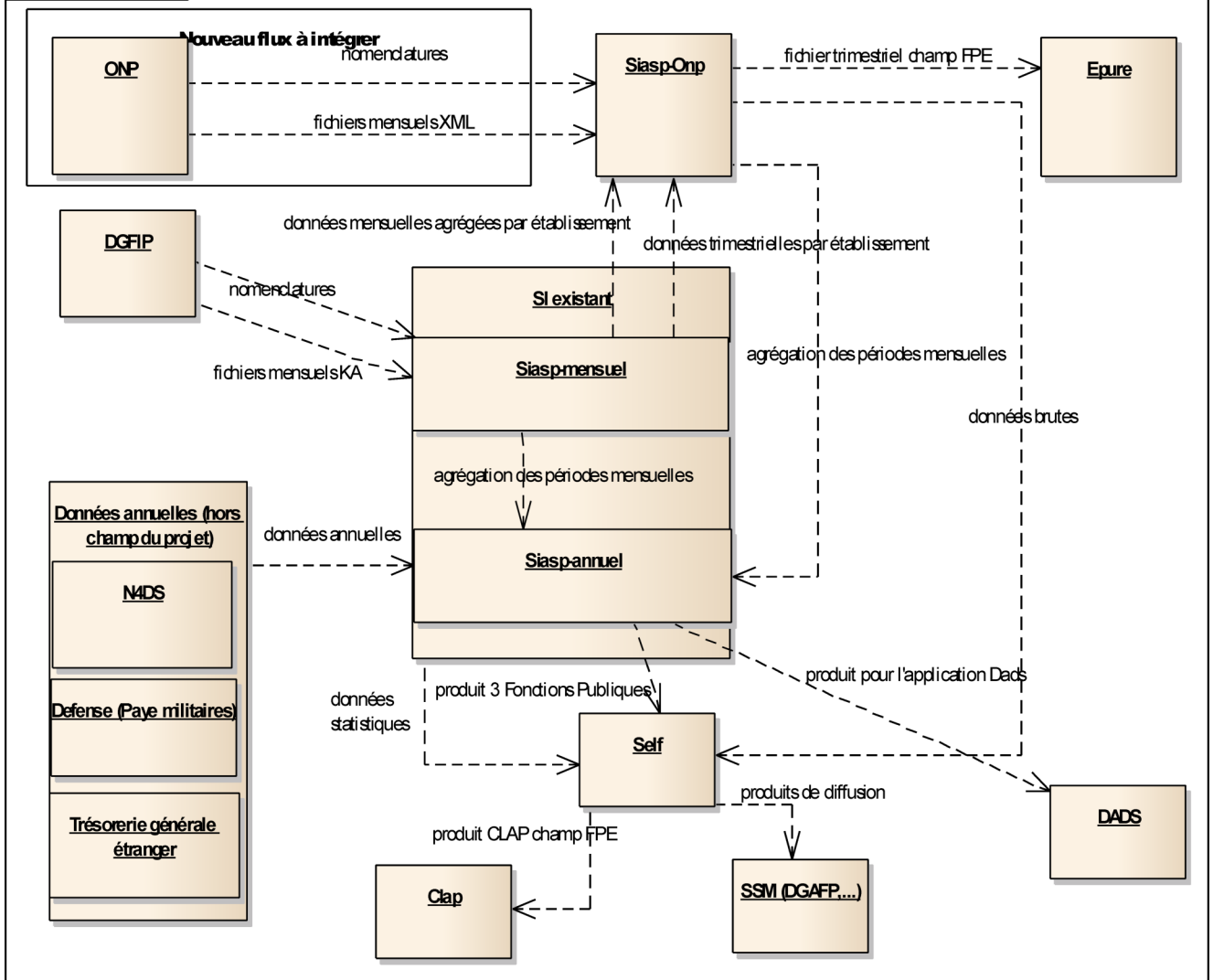
Ce lotissement ainsi que les calendriers sont étudiés dans l'étude préalable.

On distingue deux types d'échanges :

- Les flux « entrants » qui viennent de l'extérieur de l'application Siasp-ONP et qui sont originaires de l'ONP ou de l'existant Siasp ;
- Les flux « sortants » qui sont produits par le nouveau système d'information et à destination du système d'information existant ou d'un acteur hors du système ;

Le diagramme ci-dessous met en évidence les flux entre l'application Siasp-ONP et les autres applications.

objet Echange de flux



2.5.1 Synthèse des flux entrants et sortants du système

FLUX ENTRANTS :

Un ou plusieurs flux au format XML devraient nous parvenir en provenance de l'ONP. Mensuellement, les fichiers de paye des ministères raccordés à l'ONP d'un volume variant de la centaine de Mo au Go. Ce sont des données ultra-sensibles (donc d'un niveau de confidentialité très élevé). À cela s'ajouteraient des flux à périodicité encore non connue dont la taille ne devrait pas excéder la dizaine de Mo contenant les métadonnées et/ou nomenclatures nécessaires à l'exploitation des fichiers mensuels.

Il y aurait également un flux entrant ne dépassant pas la centaine de Mo qui proviendrait du SI existant pour traiter l'exhaustivité dans les phases de double alimentation.

Le dernier flux entrant vient aussi de Siasp existant et ne dépasse pas la centaine de Mo. Il consiste en un transfert des données mises au format Épure sur le champ de la DGFI.

FLUX SORTANTS :

L'application Siasp-ONP alimentera Épure par le biais d'un fichier texte, appelé fichier Pléiade, d'un volume de 40Mo.

Les données issues de Siasp-ONP devront également alimenter le Siasp existant, une contrainte la volumétrie: 1 fois par an, volumétrie à terme de l'ordre de 5Go à transférer entre les deux bases.

La MOA aimerait avoir un accès en libre service aux données brutes (fichier de production et données dans les bases de production).

2.6 Exigences et contraintes technico-fonctionnelles

Les exigences correspondent à l'ensemble des besoins exprimés par la maîtrise d'ouvrage pour garantir une application de qualité. Les contraintes sont d'ordre plus organisationnel, matériel ou environnemental. Elles seront pour partie prises en compte dans un contrat d'application lors du passage en production, une fois que des solutions techniques seront trouvées.

2.6.1 Exigences (version provisoire sur la base de l'existant)

➤ De performance et de qualité de services

Indicateurs	
Fraîcheur des données de production accessibles aux responsables d'enquête (clone de la base de données)	Entre 0,5 et 1 jour

➤ De disponibilité et de robustesse

Indicateurs	
Taux de disponibilité de l'application (hors période de mise en production)	< 6 jours d'indisponibilité par an (soient 12 heures par mois) sur le temps ouvré, soit un taux de disponibilité de 97%
Durée maximale d'indisponibilité	< 6 heures en général
Délai de détection des incidents	Délai entre le moment où l'incident est réel et le moment où celui-ci est détecté (HO / HNO). Cette exigence ne concerne que les incidents liés à l'infrastructure. Bloquant : 2 heures / 4 heures Majeur : 4 heures / 8 heures Mineur : 12 heures / 24 heures
Délai de traitements des incidents	Délai souhaité de résolution de l'incident à partir du moment où celui-ci a été signalé (incident infrastructure) ou reçu (incident applicatif). Bloquant : 12 heures Majeur : 48 heures Mineur : 5 jours
Délai de gestion du changement (changement d'infrastructure matérielle, changement de version, ...)	<u>Délai de mise en place d'un changement applicatif</u> (de la réception des sources à la mise en production complète) < 48 heures (15 jours de préavis) <u>Délai de mise en place d'un changement de données</u> (de la livraison des données à la mise en production complète) < 24 heures (15 jours de préavis) <u>Délai de mise en place d'un changement d'infrastructure et de système</u> (de la notification à la mise en place). 1 mois (sauf procédure exceptionnelle) - (3 mois de préavis)

➤ De confidentialité

Besoins	
Authentification	Utilisation du composant SSO
Habilitations	Habilitations en fonction du profil des acteurs : groupes définis dans le LDAP, utilisateurs associés à l'un des groupes
Chiffrement	

➤ D'intégrité

Besoins	
Protection des données	Haute protection
Sauvegarde	<p>Périodicité des sauvegardes applicatives (données + fichiers sources)</p> <p>Perte de travail maximale acceptée : une journée (sauvegarde quotidienne).</p>

➤ De traçabilité

Besoins	
Traces des événements	<p>Logs applicatives mises tous les jours à disposition des utilisateurs pour une durée d'un mois, conservées 1 an par le site de production.</p> <p>Logs systèmes non mises à disposition mais conservées 1 an.</p>
Archivage	

2.6.2 Les contraintes

➤ De volumétrie

Besoins	indicateurs
Stockage de données (en début de projet, cible, à moyen terme)	<p>Base de données : ? Go en cible Schéma actuel 500Go</p> <p>Système de fichiers : Hypothèse : 100 Go par an ? actuel (40Go)</p>
Échanges de données (volume des flux)	Ordre de la centaine de Mo au Go en provenance ONP ?

➤ De sécurité

<i>Besoins</i>	
<i>Respect de la sécurité des flux entre les zones internes et DMZ</i>	<i>Néant</i>
<i>Sécurité logicielle</i>	<i>Patches de sécurité</i>

➤ *D'ergonomie, d'accessibilité (lois sur l'accessibilité, ...)*

<i>Besoins</i>	
<i>Niveau d'ergonomie attendue</i>	<i>Important surtout pour la partie concernant le suivi de production</i>
<i>Niveau d'accessibilité à respecter</i>	<i>Cible niveau A du WCAG (application interne mais réalisée après 2010 et la généralisation du RGAA). Les pages présentées ont pour cible de respecter la norme XHTML 1.1 (HTML 4.0), avec un encodage en norme UTF-8</i>

➤ *De pérennité, d'évolutivité*

<i>Besoins</i>	
<i>Pérennité</i>	<i>Durée de vie envisagée pour l'application : 15-20 ans</i>
<i>Évolutivité</i>	<i>Changements les plus fréquents : les dessins des fichiers en entrée, avec des conséquences sur toute la chaîne en cas d'ajout ou de suppression d'une variable.</i>

➤ *D'exploitation, d'automatisation*

<i>Besoins</i>	
<i>Supervision des serveurs, des services</i>	
<i>Ordonnancement des batch</i>	<i>Traitements ordonnancés la nuit</i>
<i>Fenêtre de tirs de batch</i>	<i>Entre 22ht et 7h les jours de semaine Entre 22h le vendredi et 7h le lundi pour le weekend (Pas d'utilisation dans les DOM)</i>
<i>Administration des bases de données</i>	

3 Architecture applicative

L'objectif de l'architecture applicative est de répartir l'ensemble des composants logiques sur les niveaux appropriés de l'architecture. Les besoins fonctionnels sont donc associés à des technologies logicielles et des ressources matérielles (comme la base de données, des serveurs web, ...).

3.1 L'architecture applicative générale

Du point de vue utilisateur, il existe les deux profils suivants :

- Le gestionnaire ;
- L'administrateur d'application;

L'interface utilisateur permettant les traitements interactifs sera construite selon une architecture J2EE : elle sera constituée de pages HTML générées par un **serveur d'application Tomcat**, et visualisées par le **navigateur** du poste de travail de l'utilisateur. Seul un navigateur est donc nécessaire sur les postes utilisateurs. On s'assurera du bon affichage des pages HTML essentiellement sur le navigateur cible **Mozilla Firefox**.

La différenciation des deux postes de travail sera réalisée par le biais de profils d'utilisateurs, leur donnant accès à certaines fonctionnalités de l'application. Les utilisateurs et leur appartenance à un profil seront stockés au sein de l'**annuaire LDAP**. L'authentification à la connexion sera réalisée en utilisant le composant Insee SSO qui évite à l'utilisateur de ressaisir le mot de passe de sa session Windows.

L'application Siasp-ONP s'appuiera donc sur une architecture **Client/Serveur Web** basée sur la filière **Java/J2EE/PostgreSQL**.

L'architecture de Siasp-ONP respecte une organisation en couches :

- **Couche client** : affichage sur le poste client, ne nécessitant dans le cas du client-serveur Web qu'un navigateur.
- **Couche de présentation (web)** : partie serveur rendant possible l'accès à la partie cliente via le protocole HTTP. Elle permet à la fois d'afficher les données reçues par la couche de services à la partie cliente via le protocole http, et d'envoyer à la couche de services les informations relatives aux actions de l'utilisateur.
- **Couche services (service)** : ou couche métier, elle gère la logique de l'application et les traitements à effectuer sur les données, indépendamment de la provenance des données et de la façon dont elles seront affichées une fois les traitements effectués.
- **Couche modèle (model)** : elle contient les objets métiers dont la plupart sont persistés en base de données. Cette couche peut être accédée directement par les couches persistance et service, mais également par la couche présentation.
- **Couche persistance (dao)** : elle gère la persistance des objets du modèle en base de données.
- **SGBD et base de données** : elle stocke les données.

Le modèle général est donc le suivant :

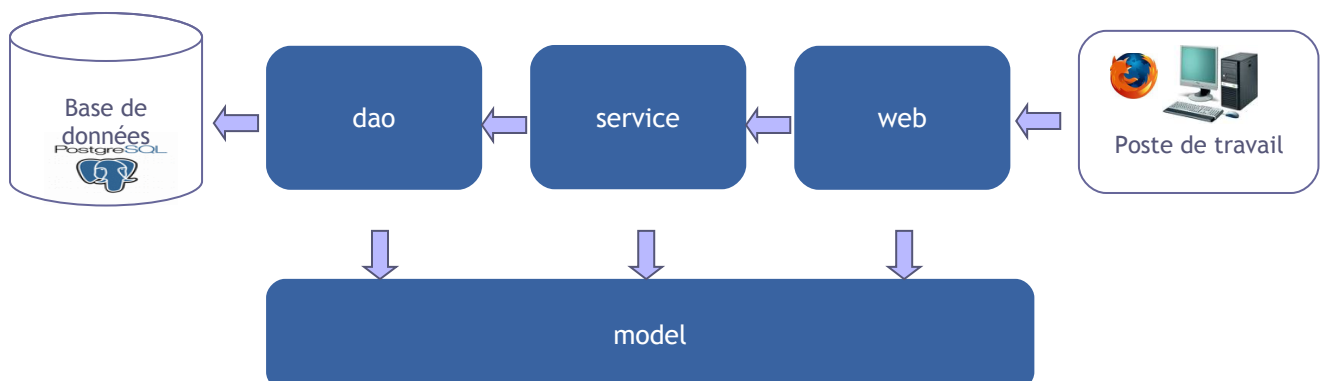


Schéma du découpage en couches

Les traitements batch s'appuieront sur un **serveur de traitement Java**. Ils seront ordonnancés (prévus pour se déclencher en fonction d'une date, de la disponibilité d'un fichier à intégrer...) pour être exécutés sans intervention des utilisateurs.

Les traitements qui sont demandés par l'utilisateur (traitements qu'il n'est donc pas possible d'ordonnancer a priori) mais sont trop lourds pour être traités en interactif feront l'objet de demandes postées. Ces traitements batch s'exécuteront en fonction de la file d'attente des autres traitements à exécuter.

Dans la cible de la nouvelle architecture de production, il est prévu que le mécanisme des demandes postées s'appuie sur des services-web.

L'ensemble des échanges avec l'ONP, qui se fait par échanges de fichiers, passera par le service d'échange de fichiers (SEF) complet (passerelle Axway et partie interne Oriade).

3.2 Accès aux données

A instruire sur la base de la note préparatoire à la troisième réunion de macroanalyse

3.2.1 Base de données de production

Les données de production seront stockées dans une **base de données PostgreSQL** organisée en trois espaces logiques :

- Un espace contenant les données brutes chargées, d'une volumétrie élevée.
- Un espace contenant les données mises au format statistique. Le volume estimé de cet espace est plus faible car il ne contient que les informations pertinentes pour le métier.
- Un espace contenant les données à destination des systèmes de diffusion (comme par exemple le produit trimestriel à destination de Épure).

Faute d'informations supplémentaires, nous ne pouvons pas être plus précis sur les volumétries de ces différents espaces

Pour effectuer leurs travaux en libre-service, les agents de la section Siasp auront besoin :

- de pouvoir extraire tout ou partie des données de la base brute contenant les variables administratives, selon des critères de sélection non figés ;
- de pouvoir accéder en lecture et extraire vers des espaces permettant une exploitation en self toutes les informations des bases de données traitant des variables statistiques, pour tracer la transformation des données depuis leur forme brute jusqu'à leur forme validée, selon des critères de sélection non prédéfinis ;
- de pouvoir extraire vers des espaces permettant de les traiter en self l'ensemble des données une fois les traitements statistiques achevés, pour les valider et préparer leur diffusion ;

Le sujet doit être instruit avec la cellule architecture.

3.2.2 Mise à disposition de fichiers

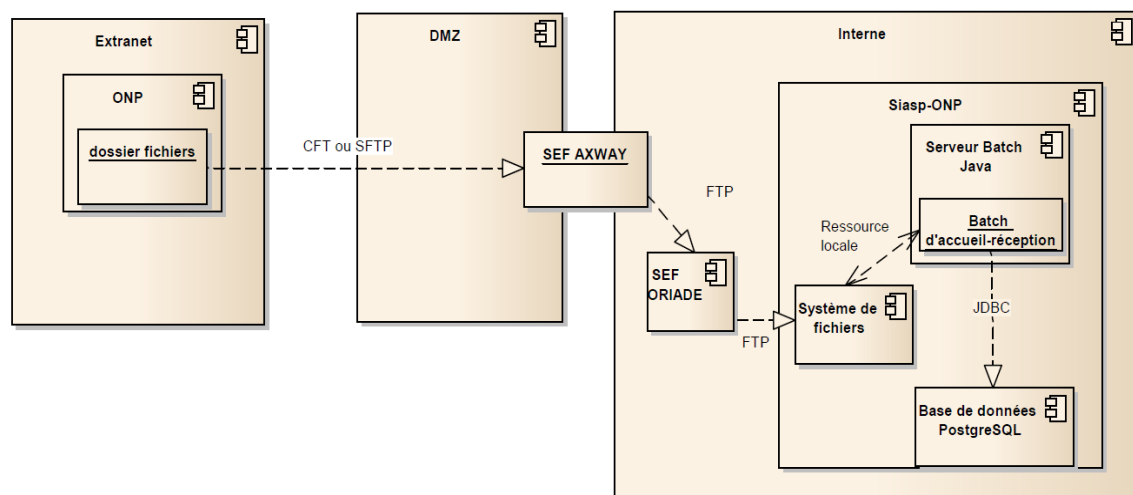
Les échanges entre les travaux libre service et la production s'effectuent par échange de fichiers, par exemple pour la mise à disposition de fichiers fournis par les utilisateurs pour leur intégration dans la base de production.

Les bilans des mises à jour des données de production par des fichiers issus de travaux selfs devront également être mis à disposition au sein d'une zone d'échange. Ils devront être conservés dans un environnement de production. L'idéal étant d'avoir une zone d'échange type FTP-appli.

Le sujet doit également être instruit avec la cellule architecture.

3.3 L'architecture applicative du module d'accueil-réception

En conséquence de l'abandon du raccordement des SIRH ministériels au SI Paye de l'ONP, le projet Siasp-ONP est suspendu. Faute d'éléments suffisants sur les autres domaines cibles du projet, l'étude de l'architecture applicative du projet Siasp-ONP est réorientée sur le domaine : Réception-contrôle-chargement des données. C'est donc l'architecture applicative de ce domaine qui sera détaillée ci-dessous.



Description du schéma d'architecture applicative :

- **Le batch** en java d'accueil réception permet d'identifier et d'introduire en base les données contenues dans les fichiers fournis par le partenaire.

Description des flux :

- L'ONP dépose les fichiers dans un dossier en FS
- Le SEF en DMZ récupère les fichiers avec le protocole sécurisé CFT ou SFTP
- Le SEF en interne transfère les fichiers au répertoire local de l'application SIASP-ONP en FTP
- Un batch Java se charge de traiter les fichiers en local (streaming java de fichier)
- Les fichiers traités sont chargés dans une base PostgreSQL en JDBC

Le batch Java permet d'enregistrer l'arrivée d'un fichier, de le contrôler puis de le charger en base.

3.4 L'organisation de l'application Java

L'application Siasp-ONP s'appuie sur un archétype Maven 2 proposé par la cellule architecture. On s'appuiera sur le Framework Spring-batch un Framework robuste permettant de développer des traitements batch.

Les tests unitaires seront automatisés au moyen des Framework **JUnit**, **DBUnit** (tests unitaires en base de données) et **Mockito** (permettant la mise en place de bouchons pour respecter le caractère unitaire des tests).

3.5 Les Framework utilisés dans le projet Java

Le projet Siasp-ONP s'appuiera, comme décrit précédemment, sur l'outil de *build* Maven2. Les Framework et autres outils utilisés dans l'application sont décrits ci-après.

3.5.1 Utilisation du Framework Spring

On s'appuiera sur le Framework Spring (Spring-core) pour les aspects suivants :

- Gestion de l'IoC (Inversion Of Control) par injection de dépendance pour obtenir une application facilement reconfigurable.
- Gestion des transactions par AOP (Aspect Oriented Programming) en utilisant le tag `@Transaction` sur les méthodes concernées.

3.5.2 Utilisation du Framework Spring-batch

On s'appuiera sur le Framework Spring-batch afin de simplifier et d'intégrer de façon modulaire la lecture, la validation et l'enregistrement en base de données des données provenant de l'ONP. Ce Framework propose notamment des outils permettant de s'adapter rapidement à des changements de formats en entrée ou en sortie (fichier plat, base de données, etc.). De plus, Spring Batch possède une librairie JAXB intégrée qui gère les fortes volumétries XML.

3.5.3 Utilisation de l'ETL Talend Open Studio

Pour les échanges entrants ou sortants entre Siasp-ONP et l'existant applicatif Siasp et/ou les travaux en libre service, l'ETL Talend pourrait être utilisé, si les temps de réponse obtenus lors du **Proof Of Concept (POC)** s'avèrent satisfaisants.

Un ETL (Extract Transform Load) facilite les trois phases d'extraction de données d'un système source, leur transformation (contrôles, transcodages) et leur chargement dans un système cible.

Cet ETL Open Source dispose d'une version monoposte gratuite Talend Open Studio. Pour des raisons de coût, c'est cette version gratuite qui sera utilisée dans Siasp-ONP. Elle permet de décrire à l'aide d'une interface graphique l'opération à réaliser, en s'appuyant sur des composants techniques disponibles, puis de générer le code Java correspondant. Elle permet de produire sur cette base des JAR qui pourront être intégrés dans l'application pour être livrés à la production avec les autres traitements développés par l'EPOI.

3.5.4 3.4.3 Autres outils

Les outils suivants seront par ailleurs utilisés :

- Framework **Log4J** pour la gestion des traces ;
- Le **composant de configuration** : gestion des paramètres de configuration d'une application Java.

Annexe 1 : Quels besoins d'accès aux données dans les applications du Siera

L'expérience acquise par le DERA sur le traitement des données administratives a montré qu'il était très difficile d'anticiper tous les chocs que peuvent subir ces données, tant dans leur format (changement de norme déclarative) que dans leur contenu (modification du comportement de déclaration de certaines variables à la suite de changements législatifs). La spécification et le développement de règles et/ou de traitements à même de gérer l'ensemble des cas possibles connus directement dans les applications seraient donc à la fois très coûteux mais également inutile. C'est pourquoi les équipes métiers auront besoin de pouvoir accéder directement, facilement et de façon souple aux données manipulées par les applications.

Les processus mis en œuvre dans le Siera peuvent schématiquement être scindés en quatre parties (cf. annexe 1) :

- l'accueil des fichiers administratifs, leur contrôle formel, le chargement en base de données transmises par le fournisseur ;
- le filtrage des données administratives et leur transformation en données statistiques brutes, c'est à dire la transformation des variables administratives vers les concepts statistiques définis ;
- les traitements statistiques qui permettent d'améliorer la donnée statistique brute via des redressements (corrections de source), des imputations (corrections de champ) ou tout autre contrôle métier spécifique comme des traitements gestionnaire ;
- la confection des produits à diffuser à partir des données statistiques finales.

Par ailleurs chaque étape du processus génère des bilans d'exécution qui permettent de faire le suivi de la production. Ces bilans permettent de disposer de comptages simples d'observations traitées, mais mettent aussi en évidence les anomalies rencontrées par l'automate.

À chaque étape, les équipes métiers peuvent avoir besoin d'accéder aux données en lecture, mais également de pouvoir les rapatrier dans des espaces à elles pourront les analyser et le cas échéant les traiter. Le mode d'accès à privilégier pour la consultation dépend de l'acteur et de la finalité de la visualisation. Par exemple, consultation des données via des IHM spécifiques développées dans le cadre des projets pour les acteurs de type « gestionnaire », et pour les besoins des reprises manuelles d'observations en erreur ; consultation des bases de données via des logiciels externes aux applications (SAS, R) pour l'administrateur de l'application, ou lorsque les données consultées ont vocation à être retraitées en masse (pour la phase de validation ou pour la gestion du secret statistique).

1°) Au niveau de la phase d'accueil-réception des données

Le besoin métier à cette étape est double :

- pouvoir accéder aux données transmises par le partenaire lorsque celles-ci ne passent pas les contrôles de forme spécifiés dans les applications ;
- pouvoir extraire tout ou partie des variables administratives des fichiers transmis par le partenaire mais non utilisées dans le processus de traitement statistique, afin de pouvoir les étudier et statuer sur la possibilité de pouvoir les y inclure ultérieurement si nécessaire (via une maintenance de l'application).

S'agissant des données transmises invalides, les équipes métiers doivent pouvoir les visualiser et le cas échéant les modifier lorsque celles-ci ne sont pas complètement conformes aux données attendues. Cet événement, bien qu'exceptionnel, peut potentiellement subvenir à chaque livraison du partenaire, laquelle peut être mensuelle, trimestrielle ou annuelle. Lorsqu'il se produit, les données en erreur seront extraites des fichiers source par les applications et chargées dans un endroit spécifique du schéma de données. La transmission de données en erreur peut advenir dans deux situations : une erreur technique dans la transmission du fichier qui conduit à perdre ou corrompre une partie de l'information ; un changement délibéré (mineur) dans le dessin de fichier dont l'Insee n'a pas été informé. Formellement il est possible dans ce cas de demander un nouvel envoi, car les transferts de fichiers entre le partenaire et

l'Insee sont réalisés sous couvert d'une convention décrivant à minima la fréquence, le format et le contenu de données transmises. Toutefois lorsqu'il y a changement dans le dessin de fichier, les délais nécessaires à la modification de la chaîne de traitement du fournisseur sont rarement compatibles avec ceux de production des indicateurs statistiques. Les équipes métier adaptent le fichier transmis en reconstituant l'ancien format, ou bien demandent une maintenance (souvent légère) sur leur chaîne d'accueil-réception. Les futures applications du Siera permettront d'extraire les fichiers ou fragments de fichier en anomalie afin qu'ils puissent être corrigés par les équipes métiers, puis réinjectés en entrée de la chaîne d'accueil-contrôle.

Le volume de données extraites devrait en général être relativement faible car seuls les enregistrements en erreur seront mis de côté. Toutefois dans le cas possible où une variable est en erreur pour l'ensemble des enregistrements du fichier (par exemple erreur de typage) tous les enregistrements du fichier seront isolés pour la variable incriminée.

Les équipes métiers doivent donc à la fois pouvoir accéder aux données en erreur extraites par les applications (dans des fichiers plats ou dans des bases de données spécifiques), mais également réinjecter les données corrigées dans les chaînes de production.

Une fois les données contrôlées, elles sont chargées dans une base de donnée dite « brute ». Cette base de donnée contient l'ensemble des informations contenues dans les fichiers source sans filtrage. Cependant toutes les variables chargées n'ont pas vocation à être utilisées dans le processus statistique, soit parce qu'elles n'apportent aucune information sur les variables statistiques d'intérêt, soit parce qu'elles n'ont pas été exploitées pour mesurer leur degré de fiabilité. Ce n'est que dans l'étape de transformation des données administratives en données statistiques que les variables sont filtrées pour ne conserver que celles qui ont un intérêt avéré pour la suite des traitements statistiques.

Les équipes statistiques souhaitent pouvoir exploiter ponctuellement en self le contenu de ces variables, sur une période suffisamment longue, afin de pouvoir déterminer si elles sont susceptibles d'apporter une information pertinente à intégrer dans le processus statistique ou non. Ce type d'étude restera ponctuel (quelques fois par an), mais devrait conduire à la consultation d'un volume de données très important (plusieurs variables sur l'ensemble des périodes disponibles, soit en régime courant de 36 mois à 5 années, selon le rythme de transmission des fichiers par les partenaires, si les durées de rétention des données dans le Siera n'évoluent pas).

2°) Au niveau de la phase de traitement statistique

Les traitements statistiques mis en œuvre dans les futures applications du Siera sont multiples et potentiellement complexes, comme par exemple le traitement des valeurs atypiques, la correction de la non-réponse partielle ou l'imputation de la non-réponse totale. Une fois ces traitements achevés, les évolutions et les caractéristiques de variables d'intérêt font l'objet d'une phase de validation qui permettra de qualifier ou pas les données. Cette phase est essentielle car elle permet de garantir la fois la qualité de la donnée statistique produite, mais également la bonne exécution de l'ensemble des traitements. Elle doit donc être menée à chaque campagne de traitement des données (une fois par mois dans le cadre de Siasp-ONP, et une fois par mois et de la DSN), avant la phase de diffusion des données. Cette analyse peut conduire à réexaminer les phases intermédiaires de calcul afin de comprendre l'évolution de la variable d'intérêt, de la valeur brute à la valeur finale. Le besoin est alors de pouvoir tracer les différentes valeurs prises par la donnée et d'en comprendre l'origine des modifications. Chaque expertise demande de mener une investigation spécifique, donc de consulter des fichiers différents.

3°) Au niveau de la phase de préparation de la diffusion

À l'issue des traitements statistiques, les bases de données servant à la diffusion doivent pouvoir être déchargées dans des espaces de travail accessibles par les équipes métiers. C'est à cette étape qu'a lieu la phase de validation des données. La validation peut prendre des formes diverses (calculs d'indicateurs, comptages spécifiques, éléments de dispersion statistique, graphiques, etc.), et évoluer au cours des campagnes en fonction des besoins et des difficultés propres à chaque exercice. Compte tenu de la nature changeante des modes de validation, ces indicateurs et les traitements qui pourraient être mis en œuvre ne peuvent pas être tous spécifiés *a priori*. En dehors de comptages simples permettant de

faire un suivi de production, le développement et le calcul d'indicateurs de qualité statistique seront donc laissés à la charge des équipes statistiques, hors application.

Au-delà de la phase de validation, la préparation de la diffusion des données, qu'elles soient agrégées ou individuelles, impose de traiter le secret statistique à la fin de chaque exercice, donc, selon les cas mensuellement, trimestriellement ou annuellement. Cette opération n'a pas vocation à être réalisée par les applications mais à l'aide d'outils spécialisés développés en self et sous la responsabilité du statisticien.

4°) Au niveau de l'analyse et de la gestion des bilans

Compte tenu de la très forte volumétrie de données traitées et du nombre d'acteurs impliqués dans le processus, les applications doivent permettre de pouvoir faire le suivi de la production.

Les applications généreront à l'issue de chaque cas d'utilisation un bilan indiquant les principaux comptages issus du traitement, ainsi que les anomalies qui auraient été détectées. Les informations sont distinguées selon leur niveau de gravité :

- les erreurs ou anomalies qui peuvent avoir une incidence sur la production ;
- les éléments de comptages généraux, qui permettent de faire un suivi global des travaux ;
- les éléments de comptages spécifiques, dont l'intérêt pour l'administrateur d'application n'est que ponctuel.

Les bilans, générés sous la forme de fichiers plats historisés et stockés sur le serveur applicatif, contiennent l'ensemble de ces informations. Les notifications d'erreurs ou d'anomalies feront l'objet d'un envoi automatique par mail à l'administrateur d'application. Les indicateurs généraux devront être consultables via l'interface de l'application. Les indicateurs spécifiques ne seront présents que dans les bilans.

Le nombre de bilans généré sera élevé, surtout dans les phases d'accueil où le chargement de chaque fichier donnera lieu à un bilan. À titre d'exemple, chaque campagne DADS traite environ 2000 fichiers. Pour une analyse efficace, l'administrateur d'application doit non seulement pouvoir les consulter mais également les rapatrier pour pouvoir les traiter en masse.

Compte tenu de ces éléments, les équipes métiers auront besoin :

- de pouvoir rapatrier puis traiter dans un espace de travail avec droits d'accès restreints⁶ les données qui n'auront pas été chargées dans l'application faute de satisfaire au dessin de fichier attendu ;
- de pouvoir déposer des fichiers corrigés, pour être chargés dans l'application en début de chaîne ;
- de pouvoir extraire tout ou partie des données de la base brute contenant les variables administratives, selon des critères de sélection non figés ;
- de pouvoir accéder en lecture et extraire vers des espaces permettant une exploitation en self toutes les informations des bases de données traitant des variables statistiques, pour tracer la transformation des données depuis leur forme brute jusqu'à leur forme validée, selon des critères de sélection non prédéfinis ;
- de pouvoir extraire vers des espaces permettant de les traiter en self l'ensemble des données une fois les traitements statistiques achevés, pour les valider et préparer leur diffusion ;
- de pouvoir consulter l'ensemble des bilans générés par l'application.

⁶ Données directement nominatives, donc forte confidentialité

ANNEXE 1 : Schéma d'un processus de traitement dans le futur Siera (hors bilans)

"Schéma de
production_V3.pdf"

ANNEXE 2 : Diagramme organisationnel des traitements : l'exemple de Siasp-ONP

"Accès aux
données_V2.pdf"