Agence ou Service : I&S

Projet : SAEM

ARCHITECTURE TECHNIQUE DU SYSTEME

Rédigé par : Benjamin CATINOT	Diffusé à : SAEM
Approuvé par : Willyam GUILLERON	

LISTE DES MODIFICATIONS DU DOCUMENT

Vers.	Date	Paragraphe	Description de la modification
00	26/03/2014		Création du document
01	12/05/2014		Intégration des protocoles d'échanges
02	22/05/2014		Intégration de tiers (horodatage, AGP et télétransmission)

SOMMAIRE

1 INTRODUCTION

Le dossier d'architecture technique du système a pour but de décrire et de définir l'architecture matérielle et logicielle d'un système d'archivage électronique mutualisé pour le Conseil Général de la Gironde, le Conseil Régional d'Aquitaine, la Communauté Urbaine de Bordeaux et la ville de Bordeaux.

2 COMPOSITION DU SYSTEME

Le système d'archivage électronique mutualisé sera composé d'un outil d'archivage électronique (As@lae) (module archivage), une GED Sas (Alfresco)(module pré-versement), un outil de communication d'archive (modules communication et recherche) et un référentiel de données (module référentiel) et un système de stockage (modules stockage et coffre-fort).

2.1 OUTIL D'ARCHIVAGE ELECTRONIQUE

As@lae, le Système d'archivage électronique de l'ADULLACT est un outil sous licence libre permettant le versement, le stockage, la gestion des données descriptives, la consultation, la communication, la destruction et la restitution d'archives.

2.2 GED SAS

La GED Sas Alfresco permet aux services versants de verser des archives, de demander l'élimination et de demander la restitution de leurs archives dans l'outil d'archivage électronique. La GED Sas est l'application centrale de versements d'archives, chaque application métiers du Système d'Information qui ont besoin d'archiver du contenu déposera le contenu dans la GED Sas.

2.3 OUTIL DE COMMUNICATION

L'outil de communication permet aux utilisateurs de pouvoir visualiser le contenu d'archives présent dans l'outil d'archivage électronique si les règles de restrictions d'accès l'autorisent.

2.4 REFERENTIEL DE DONNEES

Le référentiel est l'outil qui stocke toutes les métadonnées utiles à la pérennisation des archives. Ce référentiel est au centre du système d'information, il peut être interrogé par chaque application dans le but de faire des recherches et ajouter/modifier du contenu.

2.5 TIERS DE TELETRANSMISSION

Le tiers de télétransmission est un opérateur habilité par le Ministère de l'Intérieur de L'Outre-Mer et des Collectivités Territoriales (MIOMCT) à transporter les flux de données circulant entre la collectivité, les administrations centrales et les établissements public locaux. (Actes, Hélios, …)

2.6 TIERS D'HORODATAGE

Le tiers d'horodatage est un opérateur qui consiste à associer une date et une heure à un évènement, une information ou une donnée informatique dans le but d'enregistrer l'instant auquel une opération a été effectuée.

2.7 AUTORITE DE GESTION DE PREUVE

L'autorité de gestion de preuve porte la responsabilité de faire valoir vos documents électroniques en tant que preuve et permet de ne pas en supporter la charge.

3 ARCHITECTURE TECHNIQUE

3.1 MUTUALISATION

La mutualisation ne peut pas s'appliquer à chaque logiciel, mais elle permettra aux logiciels qui le peuvent de centraliser les données et outils pour en simplifier l'utilisation et l'accès aux données.

3.1.1 Mutualisable

L'outil de communication est mutualisable, il va permettre aux utilisateurs de pouvoir accéder aux archives de chaque collectivité via une interface unique.

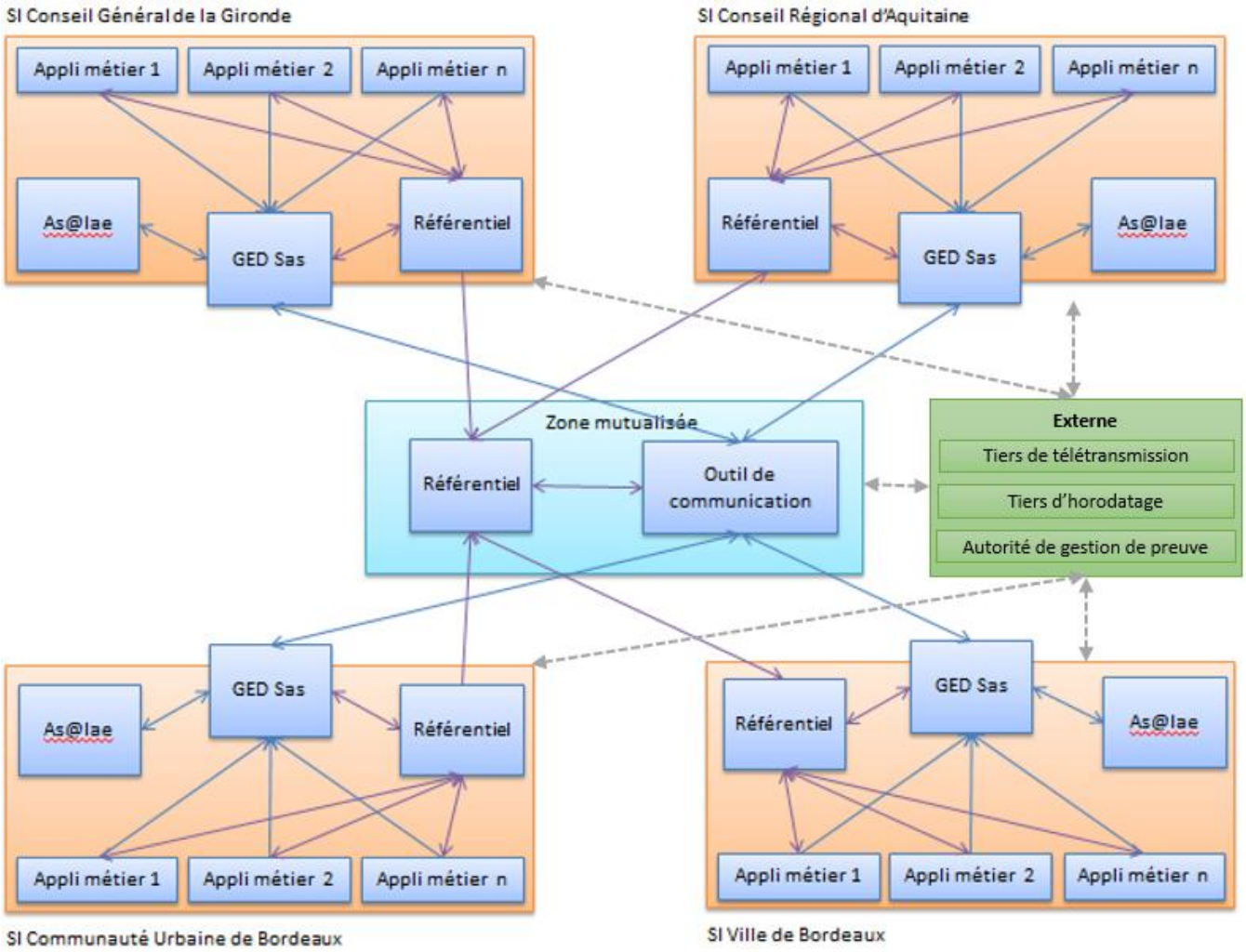
Le référentiel est mutualisable, il va permettre de rassembler en un point les métadonnées descriptives des archives de chaque collectivité. Etant donné que le référentiel est un outil pouvant être très fréquenté par les applications de chaque système d'information, il est préférable de mettre aussi une instance du référentiel dans le système d'information des collectivités. Des synchronisations ont lieu entre les référentiels des collectivités et le référentiel mutualisé afin qu'il rassemble les données.

Le tiers de télétransmission, le tiers d'horodatage et l'autorité de gestions de preuve peuvent être mutualisés car ce sont des opérateurs externes aux collectivités, habilités par des Ministères.

3.1.2 Non mutualisable

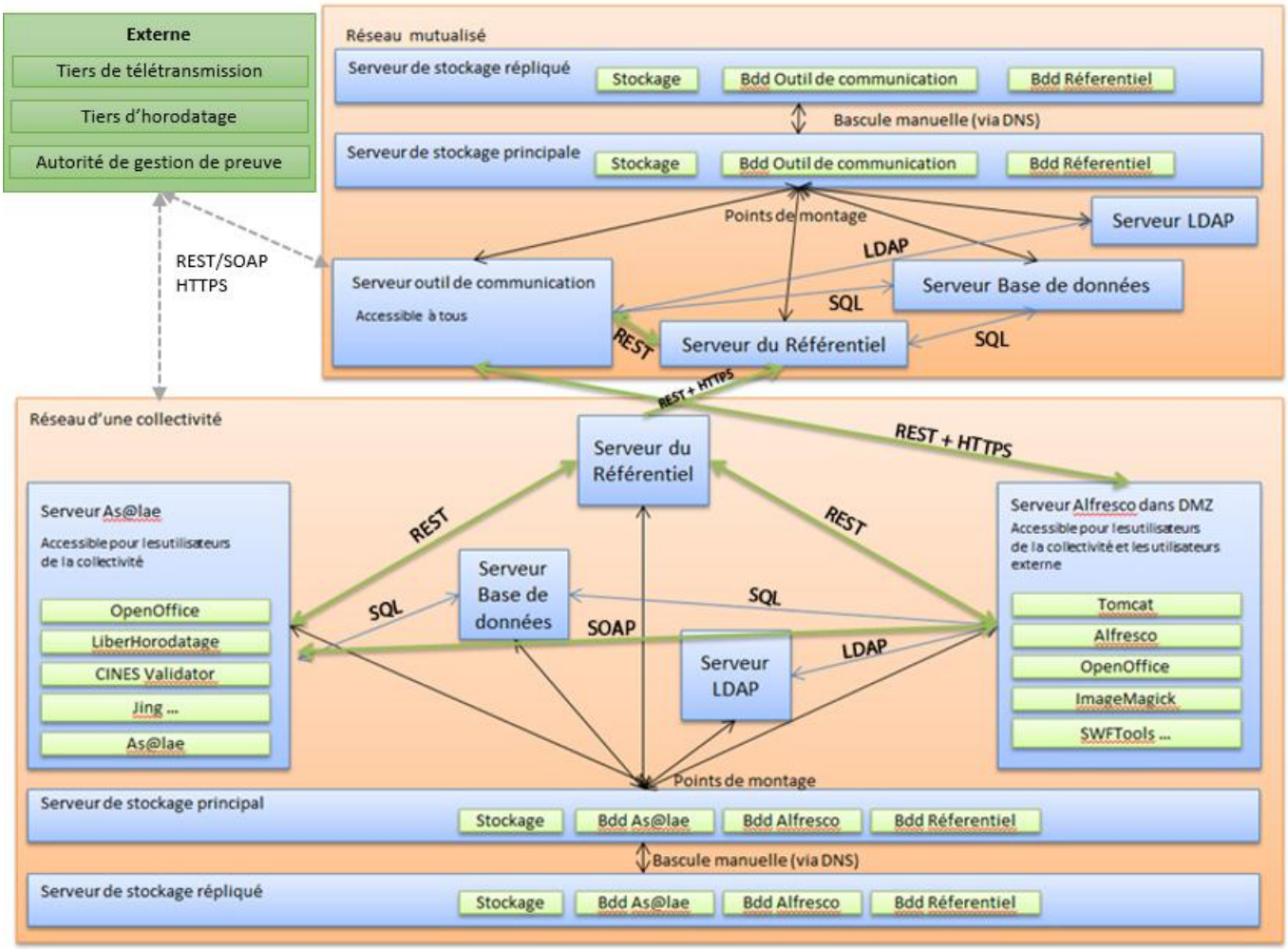
Nous avons fait le choix de ne pas mutualiser As@lae et la GED Sas pour que chaque collectivité héberge ses propres archives en toute sécurité. L'outil de communication mutualisable se servira des services de la GED Sas pour consulter ou télécharger les archives sans les stocker.

3.1.3 Schéma du système



L'As@lae présent dans le SI des collectivités permet de stocker les archives intermédiaires et définitives. Une unique instance pour les archives intermédiaires et définitives permet de limiter les configurations, les coûts de matériel, de maintenance tout en respectant le cloisonnement des données grâce au mode multi-collectivité de l'application. Ce mode permet de séparer pour les archives intermédiaires et définitives la zone de stockage et la base de donnée.

3.2 SCHEMA D'ARCHITECTURE TECHNIQUE



3.3 DESCRIPTION DE L'ARCHITECTURETECHNIQUE

3.3.1 Zones DMZ

Une zone DMZ est présente dans chaque collectivité et sur l'architecture mutualisée pour rendre accessible des applications depuis l'extérieur.

La DMZ est un réseau à part à la fois accessible depuis le réseau interne que depuis l'extérieur. Un « reverse proxy » sera présent pour régler et contrôler les accès entre l'extérieur et les serveurs de la DMZ. La politique de sécurité mise en œuvre sur la DMZ est généralement la suivante :

- Trafic du réseau externe vers la DMZ **autorisé** ;
- Trafic du réseau externe vers le réseau interne **interdit** ;
- Trafic du réseau interne vers la DMZ **autorisé** ;
- Trafic du réseau interne vers le réseau externe **autorisé** ;
- Trafic de la DMZ vers le réseau interne **interdit** ;

- Trafic de la DMZ vers le réseau externe **refusé** (sauf exception).

Dans notre cas, la DMZ pourra communiquer avec les GED Sas des collectivités et le référentiel et l'outil de communication de l'architecture mutualisé.

3.3.2 Machines virtuelles

Chaque architecture sera composée de serveur virtuel. Les avantages de la virtualisation sont :

- utilisation optimale des ressources d'un parc de machines (répartition des machines virtuelles sur les machines physiques en fonction des charges respectives),
- installation, déploiement et migration facile des machines virtuelles d'une machine physique à une autre, notamment dans le contexte d'une mise en production à partir d'un environnement de qualification ou de pré-production, livraison facilitée,
- économie sur le matériel par mutualisation (consommation électrique, entretien physique, surveillance, support, compatibilité matérielle, etc.)
- installation, tests, développements, cassage et possibilité de recommencer sans casser le système d'exploitation hôte
- sécurisation et/ou isolation d'un réseau (cassage des systèmes d'exploitation virtuels, mais pas des systèmes d'exploitation hôtes qui sont invisibles pour l'attaquant, tests d'architectures applicatives et réseau)
- isolation des différents utilisateurs simultanés d'une même machine (utilisation de type site central)
- allocation dynamique de la puissance de calcul en fonction des besoins de chaque application à un instant donné,
- diminution des risques liés au dimensionnement des serveurs lors de la définition de l'architecture d'une application, l'ajout de puissance (nouveau serveur etc) étant alors transparente.

3.3.3 As@lae

Un As@lae sera installé sur un serveur dans chaque système d'information des collectivités.

As@lae est accessible uniquement depuis le réseau interne des collectivités, et est connecté à deux bases de données (serveur de base de données) et à deux zones de stockage (serveur de stockage) pour gérer les archives intermédiaire et définitive.

Les archives intermédiaires et définitives sont donc gérées via la même application. Afin de bien cloisonner ses différentes archives, As@lae est configuré en mode multi-collectivités.

3.3.4 Outil de communication

L'outil de communication est installé uniquement sur un serveur dans le système mutualisé. L'intérêt premier est de rassembler à une unique adresse, l'accès aux archives de chaque collectivité en fonction des règles de restrictions d'accès de celles-ci.

L'outil de communication est connecté à une base de données (serveur de base de données), à une zone de stockage (serveur de stockage) et à un service d'annuaire (serveur LDAP) pour gérer ses comptes utilisateurs.

L'outil de communication doit pouvoir demander la communication d'archives aux GED Sas des collectivités, pour cela, des règles proxy sont configurées pour autoriser l'outil de communication à communiquer avec les serveurs des autres collectivités.

3.3.5 Référentiel

Le référentiel est installé sur un serveur dans le système d'information de chaque collectivité, et sur un serveur du système d'information mutualisé.

Le référentiel est connecté à une base de données (serveur de base de données) et à une zone de stockage (serveur de stockage).

Les référentiels des systèmes d'informations des collectivités doivent pouvoir se synchroniser avec le référentiel mutualisé, pour cela des règles de proxy sont configurées pour autoriser le référentiel à communiquer avec le référentiel du serveur mutualisé.

3.3.6 Alfresco

Alfresco est installé sur un serveur dans le système d'information de chaque collectivité. Ce serveur doit être disponible de l'extérieur pour permettre à l'outil de communication de demander des communications d'archives ou à des applications métiers externes pour déposer ses archives.

Alfresco est connecté à une base de données (serveur de base de données), à une zone de stockage (serveur de stockage) et à un service d'annuaire (serveur LDAP).

3.3.7 Base de données

Un serveur de base de données est disponible dans le système d'information de chaque collectivité et dans le système d'information mutualisé dans le but de gérer les données de chaque applications de leurs systèmes qui le nécessite. Stockage

Un serveur de stockage suppléant au système de fichier des machines virtuelles est utilisé. La communication avec ce système de stockage se fait par l'intermédiaire de points de montage.

Sont stockés dans ce serveur de stockage :

- Les données des bases de données
- Les fichiers des applications

Le système de stockage fait l'objet d'une sauvegarde permanente par réplication mécanique sur un deuxième serveur de stockage.

3.3.8 Annuaire

Des services d'annuaires sont présents dans les systèmes d'information des collectivités et dans l'architecture mutualisée. Les applications du système d'information se connectent à ce service d'annuaire afin d'authentifier les utilisateurs.

3.3.9 Tiers de télétransmission, Tiers d'horodatage et Autorité de gestion de preuve

Ces opérateurs externes sont équipés de service web pour permettre respectivement la télétransmission, l'horodatage et la gestion de preuve. Chacun de ces opérateurs doit respecter les recommandations du Référentiel Général de Sécurité.

4 EXIGENCES TECHNIQUES

4.1 PREREQUIS AS@LAE

Les logiciels tiers requis coté serveur pour l'installation de la solution As@lae sont les suivantes :

- Oracle Java SE Development Kit (JDK 1.6)
- PostgreSQL (version 8.2 minimum)
- OpenOffice.org (Dernière version conseillé)
- CINES Validator
- GEDOOO
- LiberHorodatage
- Serveur SMTP
- Jing
- Apache (version 2) avec le module php-soap php5-pgsql php-xsl php-curl
- Clamav

4.2 PREREQUIS ALFRESCO

Les logiciels tiers requis coté serveur pour l'installation de la solution Alfresco sont les suivantes :

- Oracle Java SE Development Kit (JDK 1.7)
- PostgreSQL (version 8.2 minimum)
- OpenOffice.org (Dernière version conseillé)
- ImageMagick
- SWF Tools (Version 0.8.1 minimum)
- Serveur SMTP