

LDD_P11_Dossier d'Architecture Technique_V2.1

ASP DATALAKE

Exported on 2024-12-18 12:09:01

Table of Contents

1.1 Objectifs.....	8
1.2 Périmètre.....	8
1.3 Localisation.....	8
2.1 Conventions de nommage	9
3.1 Réseau	10
Présentation générale.....	10
3.1.2 Réseau interne.....	11
3.1.3 Accès RIE	13
3.1.4 Accès Internet	14
3.1.5 Accès Capgemini	15
3.2 Schéma de l'architecture.....	15
3.3 Listes des environnements.....	16
3.3.1 Machines physiques du stockage à froid	16
3.3.2 Machines virtuelles	17
3.3.3 Points de montage.....	17
3.3.4 Plateforme d'échanges (PFE) de fichiers.....	17
3.4 Bases de données.....	18
3.5 Composants et services	18
3.5.1 Load-Balancing.....	18
3.5.2 Configuration homogène - Ansible.....	19
3.5.3 Sauvegarde et réplication VEEAM.....	19
3.5.4 Concentration des logs système - GrayLog	20
3.5.5 Supervision Zabbix	21
3.5.6 Antivirus / Antimalware.....	21
3.5.7 Bastion - CyberARK.....	21
3.5.8 LDAP - ForgeRock Directory Server	22
3.5.9 Service NTP.....	22
3.5.10 DNS	22
3.5.11 Gestion des mises à jour.....	23
3.5.12 Relais de mails.....	24
3.5.13 Spécifiques à l'application.....	24
3.5.14 Environnement physique du Datalake	24
4.1 Schéma d'architecture applicative.....	26
4.1.1 Schéma simplifié.....	26
4.1.2 Exemple d'articulation entre les composants applicatifs	26
4.1.3 Schéma d'architecture	27
4.2 Contexte du projet	30
4.2.1 Enjeux métiers	30
4.3 Exigences fonctionnelles.....	30
4.3.1 Exigences fonctionnelles.....	30
4.4 Exigences non fonctionnelles.....	30
4.4.1 Contraintes existantes	30
4.4.2 Volumétrie Réplica ISIS	30
4.4.3 Objectif de sécurité	31
4.5 Description des composants	32
4.5.1 Cluster BigData.....	32
4.5.2 Conteneurisation.....	34
4.5.3 Sécurité.....	35
4.5.4 Audit.....	37
4.5.5 Base de données PostgreSQL.....	38
4.5.6 Moteur de calcul spatial PostGis.....	38
4.5.7 Plateforme d'intégration continue (PIC)	38
4.5.8 Dataiku.....	38
4.5.9 RStudio.....	40
4.5.10 MicroStrategy.....	40
4.5.12 Ordonnancement VTOM.....	41
4.5.13 Serveurs sur lesquels est installé VTOM	42
4.5.14 Stockage Froid.....	42
6.1 Supervision.....	43

6.2 Sauvegardes	43
6.3 Ordonnancement.....	43
6.4 Maintenance	43
7.1 Capacités d'accroissement	44
7.2 Sécurité	44
7.3 Mise à disposition de logs système.....	44
8.1 Disponibilité du service.....	45
9.1 Liste des VLAN / subnet.....	46
9.2 Liste des adresses mises à disposition	46
3 Liste des VM.....	46
9.4 Liste des points de montage	49

Introduction

Terminologie

Abréviations ou acronymes

Abréviation ou acronyme	Signification
ASP	Agence de Service et de Paiement
BDD	Base De Données
CCTP	Cahier des Clauses Techniques Particulières
GTR	Garantie de Temps de Rétablissement
LLD	Low-Level Design – Document d’implémentation d’un composant
NAS	Network-Attached Storage (Réseau de stockage – mode fichier)
RIE	Réseau Inter-Etat (Réseau privé de l’administration)
PRA	Plan de Reprise d'Activité
RAC	Real Application Clusters. RAC Oracle
SAN	Storage Area Network (Réseau de stockage – mode bloc)
VPN	Virtual Private Network : réseau privé virtuel
DDOS	Distributed Denial Of Service
SLA	Service Level Agreement – Taux de disponibilité
RTO	Recovery Time objective / Durée d’indisponibilité Maximale
RPO / PDMA	Recovery Point Objective / Perte de données maximale admissible
BLOB	Binary Large Object Type de donnée permettant le stockage de données binaires (des fichiers de type image, son ou vidéo) dans le champ d'une table d'une base de données.
DAS	Terme utilisé pour un système de disque en attachement direct, par opposition au SAN ou NAS qui sont en attachement réseau. Le système disque ainsi installé n'est accessible directement qu'aux ordinateurs auquel il est raccordé.
HDFS	Hadoop Distributed File System Système de fichiers distribué conçu pour stocker de très gros volumes de données sur un grand nombre de machines équipées de disques durs banalisés.
JSON	JavaScript Object Notation Format de données textuelles dérivé de la notation des objets du langage JavaScript permettant de représenter de l'information structurée comme le permet XML par exemple.
JBOD	JBOD signifie « Just a Bunch Of Disks » (littéralement : « juste un paquet de disques »). Il s'agit d'une mise à disposition de disques durs, peu importe leur taille, sans configuration RAID. En cas de panne d'un des disques, les données situées sur ce dernier sont perdues, mais celles des autres disques restent accessibles. Cette architecture disque est recommandée dans les clusters Hadoop car les

Abréviation ou acronyme	Signification
	mécanismes de réplication sont gérés par le Framework lui-même.
KMS	Key Management System Service qui permet de créer et de contrôler facilement les clés de chiffrement utilisées pour chiffrer des données.
XML	eXtensible Markup Language Métalangage informatique de balisage permettant de créer des documents autodescriptifs.
ZE	Sas interne et externe à la DNSCE. Il s'agit d'un stockage temporaire. Les applications, qui tirent la demande (pas de mode push depuis l'extérieur) scrutent la Zone d'échange à la recherche de nouveaux fichiers.
REST	REST (REpresentational State Transfer) est un style d'architecture pour les systèmes hypermédia distribués et une approche de communication qui est souvent utilisée dans le développement de services Web. Une API compatible REST, ou « RESTful », est une interface de programmation d'application qui fait appel à des requêtes HTTP pour obtenir (GET), placer (PUT), publier (POST) et supprimer (DELETE) des données.

Définitions

Définition	Signification
Big Data	Ensembles de données dont le volume les rend difficiles à travailler avec des outils classiques de gestion de bases de données ou de gestion de l'information.
Chiffrement	Le chiffrement est un procédé de cryptographie grâce auquel on souhaite rendre la compréhension d'un document impossible à toute personne qui n'a pas la clé de (dé)chiffrement. Ce principe est généralement lié au principe d'accès conditionnel. Dans le cas d'un cluster Hadoop, on peut distinguer le chiffrement des données stockées (ex : HDFS encryption) du chiffrement des données avant émission sur le réseau.
Data Lake	Un lac de données (Data Lake) est un référentiel de stockage qui conserve une grande quantité de données brutes dans leur format d'origine. Le lac produit aussi des données dans les espaces d'analyse dans un format associé à chaque cas d'usage.
Data mining	Exploration de données. Extraction d'un savoir ou d'une connaissance à partir de grandes quantités de données, par des méthodes automatiques ou semi-automatiques en utilisant un ensemble d'algorithmes issus de disciplines scientifiques diverses telles que les statistiques, l'intelligence artificielle ou l'informatique, pour construire des modèles à partir des données, c'est-à-dire trouver des motifs, et d'en extraire un maximum de connaissances. L'utilisation de ce savoir permet de résoudre des problèmes divers tels que la gestion de la relation client, la maintenance préventive, la détection de fraudes, l'optimisation de sites web.
Data scientist/ Data miner	Le Data scientist est un expert de l'analyse pointue de données massives ("big data"). Il crée, à partir de sources de données multiples et dispersées, des modèles exploratoires, prédictifs et prescriptifs dont les résultats permettent la mise en place d'une stratégie efficace d'actions métiers répondant à une problématique. Exemple : Modélisation de cas

Définition	Signification
	<p>frauduleux/violation des règles qui permettent d'enrichir la base RMS avec des règles plus complexes, intégrant plus de variables et leurs interactions qui, par la suite, contribuent à une gestion plus efficace des vérifications.</p> <p>Il est donc spécialisé en statistiques, algorithmie complexe de type machine learning. optimisation et recherche opérationnelles ainsi que calculs distribués.</p>
DataNode	Élément d'un cluster Hadoop, le DataNode est le composant distribué de stockage des données dans le système de fichier Hadoop (ex : HDFS).
Développeur	Le développeur conçoit, développe et intègre de façon industrielle les flux d'alimentation de données. Il est également responsable de mettre en œuvre les transformations et enrichissements de données nécessaires au succès des algorithmes construits par le Data scientist.
Distribution	Ensemble cohérent et packagé de solutions logicielles ayant pour but la constitution d'une plate-forme Big Data. Exemples : Cloudera.
Hadoop	<p>Framework libre et open source Java destiné à faciliter la création d'applications distribuées (au niveau du stockage des données et de leurs traitements) et échelonnables (scalables) permettant aux applications de travailler avec des milliers de nœuds et des pétaoctets de données. Ainsi chaque nœud est un serveur de commodités.</p> <p>L'ensemble des composants du Framework Hadoop est conçu dans une logique distribuée et de haute disponibilité permettant l'absorption des pannes matérielles avec aucune ou de faibles perturbations.</p>
Machine Learning	<p>Apprentissage Automatique</p> <p>Ensemble de méthodes statistiques ou issues de l'intelligence artificielle, permettant aux ordinateurs d'apprendre à partir des données qui leurs sont soumises, et plus seulement d'exécuter des algorithmes.</p>
Master Node	Dans un environnement distribué Hadoop, serveur contenant les processus coordinateurs des programmes Hadoop s'exécutant sur le cluster. Il s'agit, entre autres, du ResourceManager (pour YARN – gestionnaire de ressources) et du NameNode (pour HDFS – système de fichiers distribués).
NameNode	Utilisé dans le cadre Hadoop, ce composant gère l'espace de noms, l'arborescence du système de fichiers et les métadonnées des fichiers et des répertoires. Il centralise la localisation des blocs de données répartis dans le cluster. Il est unique mais dispose d'une instance inactive, appelé Standby NameNode. Ce Standby NameNode gère la continuité du fonctionnement du cluster Hadoop en cas de panne du NameNode d'origine.
Objet métier	Structuration standardisée de données produites par un téléservice douanier et mise à la disposition d'applications consommatrices
Parquet	Apache Parquet est un format de données libre et open source orientée colonnes de l'écosystème Apache Hadoop. Il est similaire aux autres formats de fichiers de stockage en colonnes disponibles dans Hadoop, à savoir RCFile et Optimized RCFile. Il est compatible avec la plupart des frameworks de traitement de données dans l'environnement Hadoop (Spark, MapReduce). Il fournit des schémas de compression et d'encodage de données efficaces avec des performances améliorées pour traiter des données complexes en masse.

Définition	Signification
POC	Proof Of Concept Démonstration de faisabilité d'une solution informatique sous la forme d'une réalisation incomplète préalable à l'élaboration d'un prototype pleinement fonctionnel.
Raffinée	Désigne une information produite par un traitement réalisé sur d'autres données (brutes ou déjà raffinées).
RessourceManager	Composant maître du gestionnaire de ressources de Hadoop – YARN. Il coordonne l'allocation des ressources du cluster (CPU / RAM). Afin d'assurer une continuité de service un Standby NameNode peut être mise en place en cas de défaillance du service d'origine.
Scalabilité	Capacité d'une architecture à évoluer en cas de montée en charge si nécessaire. On distingue la scalabilité horizontale et la scalabilité verticale.
Scalabilité horizontale	Possibilité d'ajouter des serveurs d'un type donné. Par exemple : ajout possible de serveurs DataNode pour augmenter la surface de stockage et la puissance d'un cluster Hadoop.
Scalabilité verticale	Possibilité d'upgrader un serveur (ajout de processeurs, RAM, disques...)
Service Node	Dans un environnement distribué Hadoop, serveur contenant les processus portant la couche de présentation des données et offrant des services aux utilisateurs finaux (ex : Hive, Client Spark, Zeppelin,...).
Shared Nothing Architecture	Architecture de calcul distribué dans lequel chaque nœud est indépendant et autonome, et il n'y a pas un seul point de contention au sein du système. Plus précisément, aucun des nœuds ne partagent les ressources CPU, mémoire ou de stockage.
Slave Node	Dans un environnement distribué Hadoop, serveur responsable de l'exécution des traitements sur la partie des données qui lui est affectée.
Structurée/Non structurée (donnée)	Les données structurées sont celles dont l'ensemble des valeurs possibles est déterminé et connu à l'avance. Les données non structurées sont potentiellement différentes et difficiles à catégoriser a priori. De façon générale, les données non structurées sont des données textuelles (documents hors formulaires, mails, réponses libres à des questions, etc.).
Virtualisation	La virtualisation est le processus consistant à créer une version virtuelle d'une entité physique. La virtualisation peut s'appliquer à des ordinateurs, à des systèmes d'exploitation, à des périphériques de stockage, à des applications ou à des réseaux.

1. Contexte

1.1 Objectifs

Ce document présente l'architecture technique de l'infrastructure et l'architecture applicative du projet ASP Datalake.

Il décrit les éléments qui la composent (matériel, logiciel, réseau), leur mise en œuvre et les relations entre ces éléments. Il détaille également les mesures de sécurité utilisées pour protéger les données et les accès.

Il comprend aussi l'architecture applicative mise en place pour le projet ASP Datalake.

Il s'appuie sur les exigences détaillées dans le CCP **[DR][2]**

1.2 Périmètre

Les équipements composant la plateforme Datalake ci-dessous sont décrits dans ce document :

- Production
- Pré-production et DataLab
- Dev/Rec/Int

Les infrastructures et équipements qui composent les services mutualisés, comme la supervision, la plateforme d'intégration continue ou la sauvegarde par exemple, seront détaillés dans des documents distincts. Ces services mutualisés sont les outils d'administration de Capgemini qui ne sont pas directement installés dans la zone réseau OVH de l'ASP. Cf DAT relatif au marché subséquent n°1 **[DR][10]**.

(NB : aucun document technique ARCGIS n'a été identifié sur le périmètre ISIS).

1.3 Localisation

La plateforme est hébergée chez OVH, sur les sites de Roubaix (site principal) et Strasbourg (secours). Seuls les personnels accrédités de OVH et ses mainteneurs peuvent accéder aux machines physiques.

Matériel / Virtualisation

Se référer au paragraphe §2.5 'Matériel / Virtualisation' du document **[DA][10]**.

2. Nomenclature

2.1 Conventions de nommage

Se référer au paragraphe §3 'Nomenclature' du document **[DA][10]**.

3. Architecture technique

3.1 Réseau

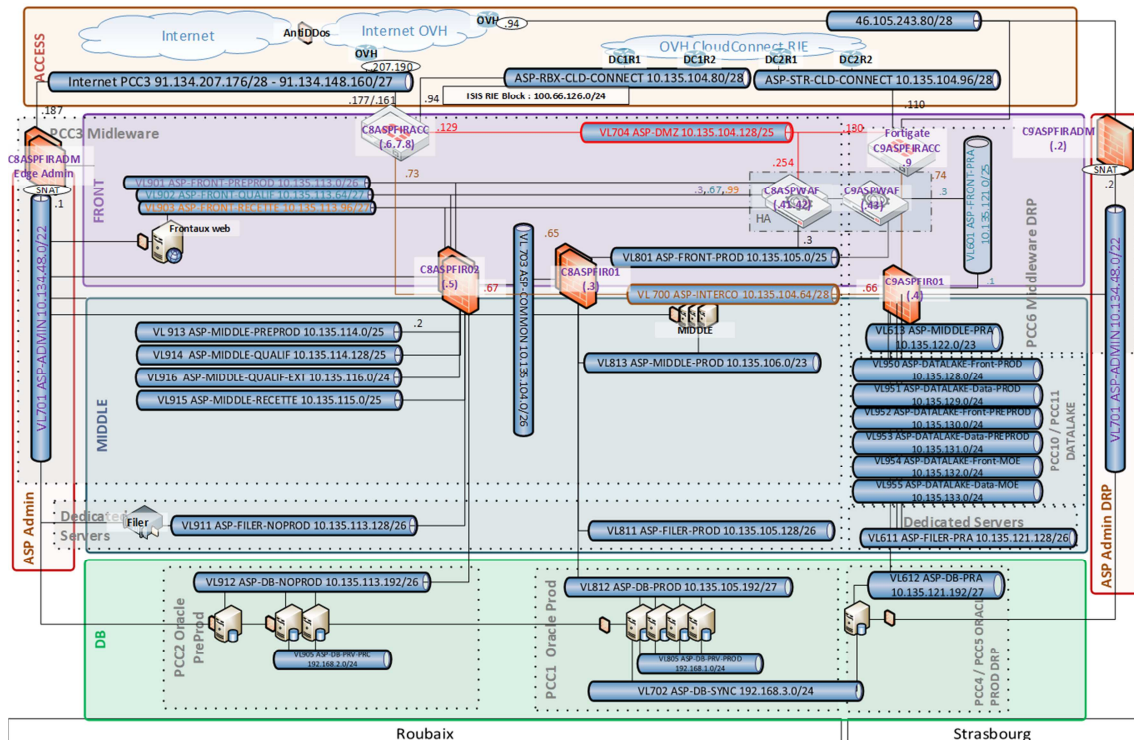
Se référer au paragraphe §4.2 'Réseau' du document [DA][10].

Présentation générale

L'infrastructure réseau mise en place doit permettre :

- La communication internet à la plateforme :
 - La communication réseau entre des VM et serveurs physiques d'un même réseau (commutation)
 - La communication réseau entre des VM et serveurs physiques de réseaux différents (routage)
 - L'isolation et le filtrage des flux
 - Isolation des flux d'administration, de hors production et de production
 - Limitation des flux entre les différentes zones de production (ISIS / RIE / Internet) aux flux strictement nécessaires
 - Le filtrage applicatif et les load balancing des flux des sites web
- L'interconnexion avec différents réseaux externes :
 - Internet pour la publication des services et l'accès à des ressources externes
 - Le réseau interministériel de l'état (RIE) pour la publication des services et l'accès à des ressources externes
 - Capgemini pour l'administration de l'infrastructure et l'accès des développeurs

3.1.2 Réseau interne



Niveau 2

3.1.2.1.a Principes

La bulle de service ASP ISIS est composée de

- 5 "private cloud" OVH (ou PCC) pour ISIS (3 à Roubaix, 2 à Strasbourg)
- Plusieurs "dedicated servers" OVH pour les serveurs de sauvegardes et les baies Netapp
- 1 "private cloud" avec stockage vSAN sur Strasbourg pour Oracle (réplicas ISIS et bases Datalake)
- 1 "private cloud" avec stockage vSAN sur Strasbourg pour Datalake (Cloudera)
- 1 "private cloud" sur Strasbourg pour Datalake Appli (à fusionner avec le PCC PRA Middleware ISIS)

Ces services sont positionnés sur le même vRack OVH. Un vRack OVH est un réseau virtuel étendu dédié permettant à tous les équipements ASP hébergés chez OVH de communiquer entre eux, dans les 2 datacenters. Grâce à cela, 2 ports groups configurés avec le même vlan ID sur chaque datacenter seront sur le même domaine de broadcast.

Les infrastructures réseaux de production et pré production (Fortinet et NSX) sont installées sur le private Cloud PCC3. Les infrastructures réseaux de PRA sont installées sur le private Cloud PCC6.

3.1.2.1.b Micro segmentation / distributed firewall

L'offre VMware inclut une fonctionnalité de micro segmentation qui permet de gérer un firewall de niveau 2 distribué sur chaque interface réseau des VM gérées par le vCenter. Cette micro segmentation est uniquement utilisée pour empêcher les rebonds par les interfaces d'administration.

3.1.2.1.c Liste des VLAN / subnet

La liste des VLAN et subnets a été déplacée dans le paragraphe Annexes.

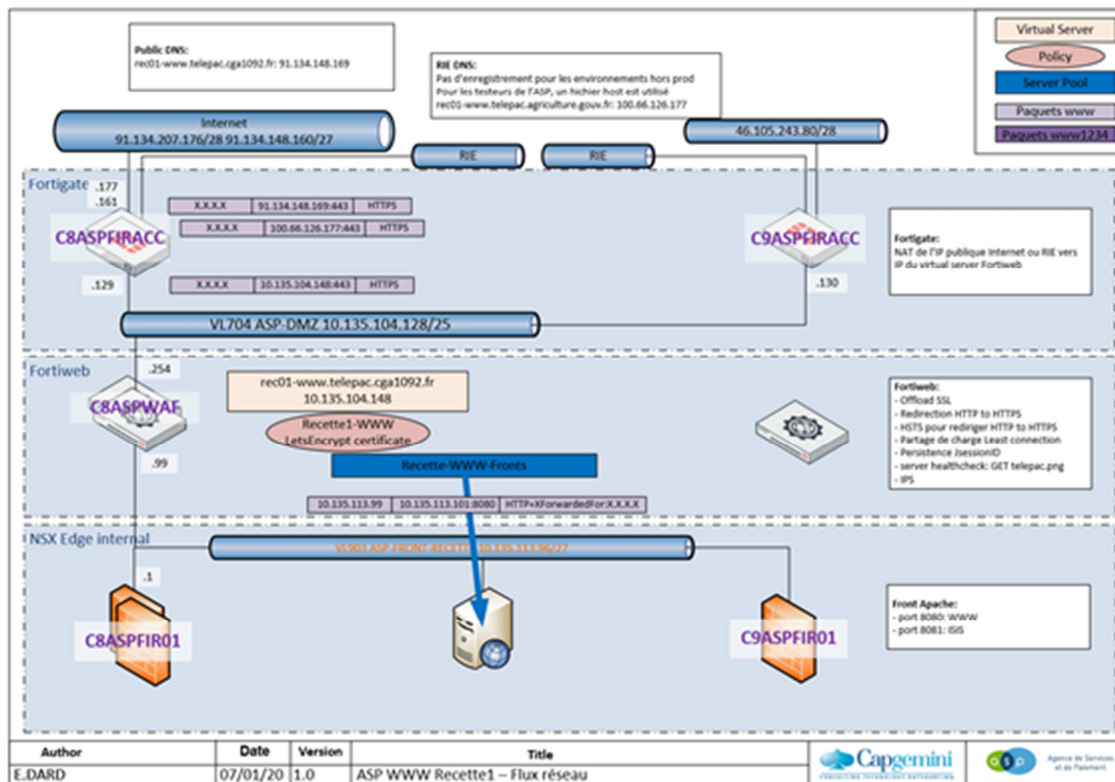
3.1.2.2 Réseau d'accès internet / RIE

Une zone d'accès, permet de centraliser les accès externes de production et de hors production.

Cette zone est constituée de :

- Un cluster firewall d'accès Fortigate connecté à internet et au RIE :
- Effectue un premier niveau de filtrage et de translation d'adresse pour les services de production et de pré production (firewall & IPS)
- En fonction de la source, il dirige les services entrants, soit sur une analyse applicative http (Fortiweb), soit directement sur les frontaux web à travers un firewall NSX Edge.
- Un cluster d'appliance de sécurité Fortiweb, qui effectue :
 - Le filtrage applicatif des flux en provenance d'internet
 - Le déchiffrement SSL
 - Le load balancing

Architecture type d'un flux web entrant traité par le Fortigate et Fortiweb :



3.1.2.3 Réseau de production / hors production

Côté production, l'architecture réseau et sécurité retenue est basée sur :

- Un cluster de firewall de second niveau NSX Edge dédié à l'interconnexion et au filtrage des réseaux internes de production
- Un cluster firewall de second niveau NSX Edge dédié à l'interconnexion et au filtrage des réseaux internes hors production

Le fait d'avoir des vlan et des firewalls dédiés aux environnements production et hors production, garantit l'étanchéité des flux entre les 2 périmètres.

L'ensemble des équipements de l'infrastructure réseau et sécurité de production et hors production est hébergé sur le private cloud Middleware de Roubaix.

On retrouve sur le private cloud de Strasbourg certains composants pré-positionnés pour un PRA :

- Un firewall Fortigate d'accès internet / RIE
- Une appliance Fortiweb
- Un firewall de second niveau NSX Edge

3.1.2.4 Réseau d'administration

Côté administration l'architecture retenue est basée sur :

- L'utilisation d'une interface dédiée à l'administration sur chaque serveur, en plus de l'interface de production qui porte la route par défaut
- L'utilisation d'un cluster de firewall NSX Edge dédié à l'administration
- Afin d'améliorer l'isolation et d'éviter les routes statiques sur le serveur, du source NAT est utilisé pour les connexions Cap vers ASP
- Pour que les serveurs puissent accéder à l'outillage Capgemini via leur interface d'administration, on fait destination NAT sur les réseaux d'administration.
- La micro-segmentation NSX pour permettre l'isolation des interfaces d'administration de chaque serveur

Il n'y a donc aucun routage activé entre la zone d'administration et la zone de production. De plus même s'il y en avait, le plan IP de production ASP n'est pas routé chez Capgemini et le plan IP de Capgemini n'est pas routé côté ASP.

Ce cluster de firewall d'administration est installé sur le private cloud Middleware de Roubaix.

Un firewall d'administration est pré positionné sur le private cloud de Strasbourg, pour être utilisé en cas de PRA.

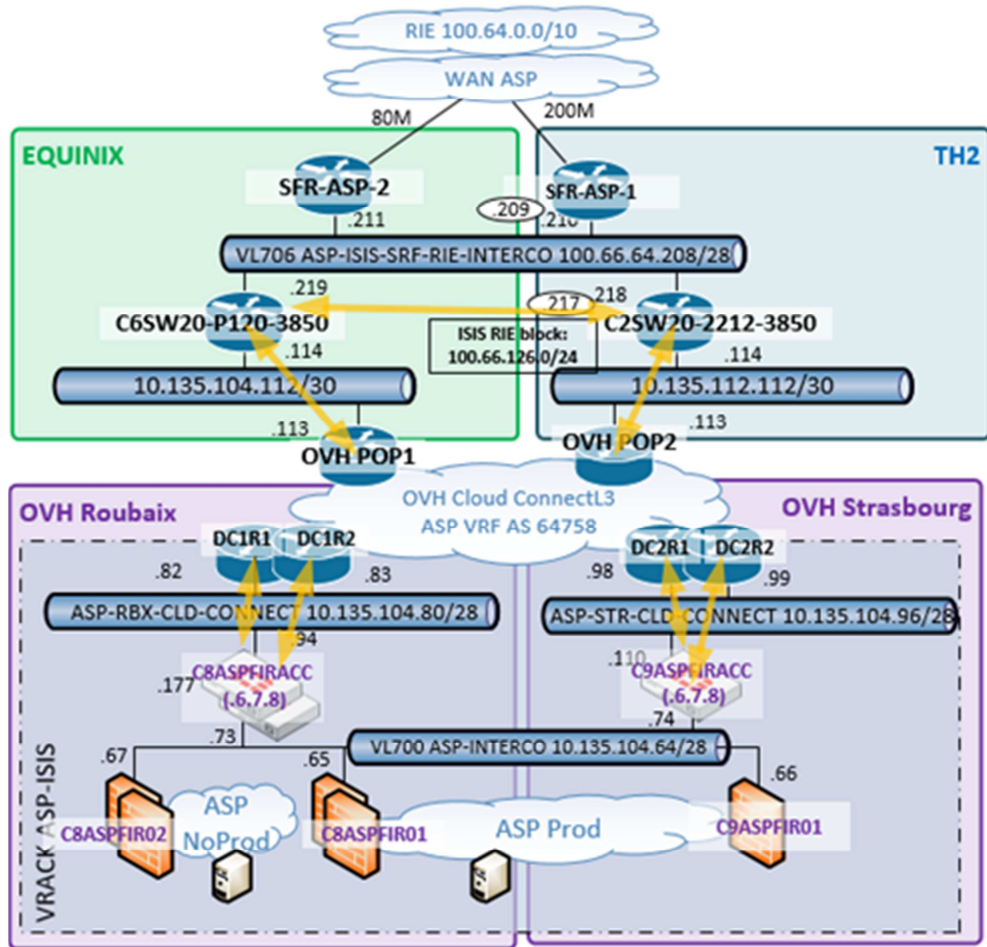
3.1.3 Accès RIE

Différents services de la plateforme utilisent l'interconnexion au WAN ASP/RIE :

- Le portail public WWW
- Le portail de gestion des dossiers ISIS
- La résolution DNS du sous domaine telepac.agriculture.gouv.fr pour le proxy interne du ministère de l'agriculture
- Mettre à disposition un partage SFTP et de serveurs Swift MQ
- Transmettre les données de paiement à l'ASP
- Transférer les fichiers photos

Les datacenters d'OVH ne permettent pas d'héberger d'équipement tiers. La solution d'accès au RIE est donc basée sur :

- Un accès au RIE en mode actif standby fourni par SFR sur les sites de Equinix et de Telehouse2 (hébergement des équipements SFR dans des baies Capgemini/Prosodie)
- Un service d'interconnexion OVH Cloud CONNECT L3 pour relier Equinix et Telehouse2 aux datacenters d'OVH
- Une paire de switches d'interconnexion RIE/OVH sur Equinix et Telehouse2
 - Une VRF de ce switch est dédiée à l'interconnexion RIE
 - Le protocole BGP côté OVH, et HSRP côté RIE, permet d'assurer la redondance des accès



3.1.4 Accès Internet

Différents services de la plateforme doivent être accessibles depuis Internet :

- Le portail public du Datalake <https://lda.asp-public.fr> à condition que les adresses IP de l'entité soient déclarées
- La résolution DNS publique du sous domaine telepac.agriculture.gouv.fr
- Mettre à disposition de partenaires, un partage SFTP

De plus, la plateforme utilise internet pour différents besoins :

- Envoi des mails aux utilisateurs

L'accès internet de la plateforme est celui fourni par OVH avec son private Cloud. Il dispose par défaut, pour chaque private Cloud, de 16 IP publiques et d'une bande passante de 1,5Gbit/s.

Les IP publiques d'OVH sont associées à un private cloud, et donc à un site physique, il n'y a pas de portabilité des adresses possible entre les sites de Roubaix et de Strasbourg.

La plateforme de production/pré production utilise donc l'accès internet OVH fourni avec le PCC3. La plateforme de PRA utilise donc l'accès internet OVH fourni avec le PCC6.

L'offre d'accès internet d'OVH intègre par défaut un filtrage entre Internet et le backbone Internet d'OVH. Par défaut, ce filtrage inclut un service anti-DDOS. Pour plus de détails : <https://www.ovh.com/fr/anti-ddos/>

3.1.5 Accès Capgemini

Capgemini possède une bulle de service hébergée chez OVH.

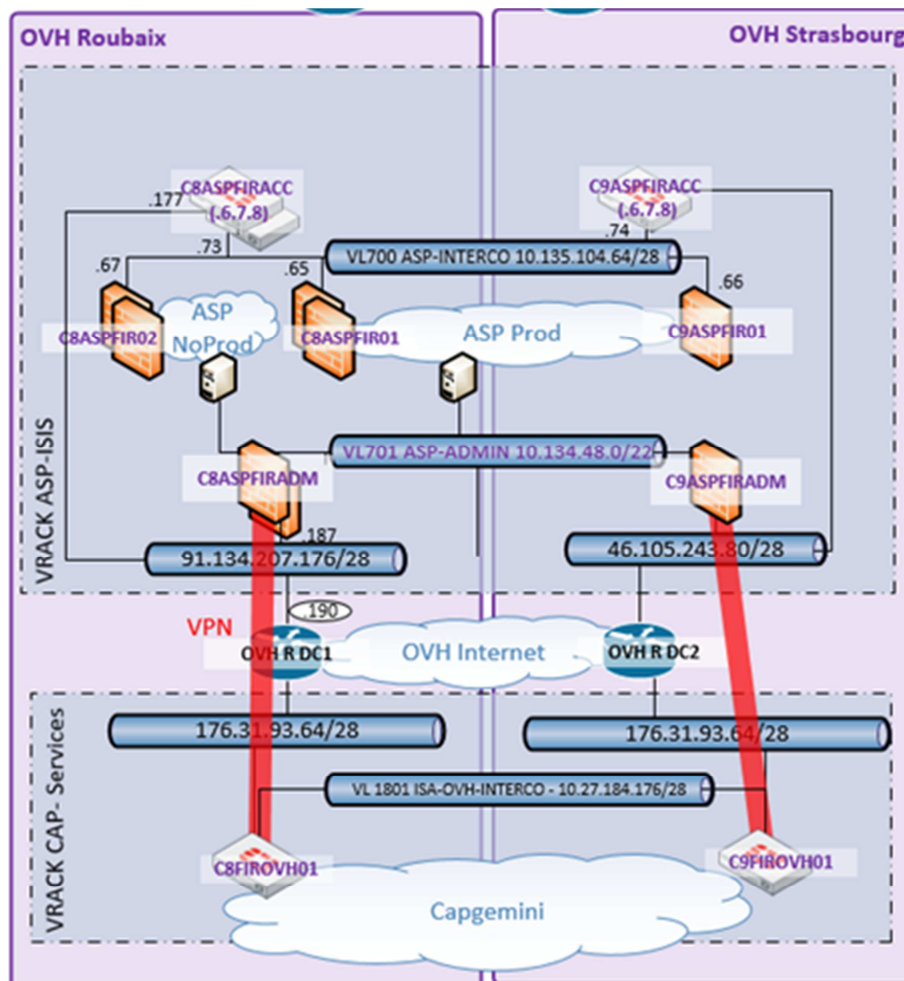
Cependant, les bulles OVH de ASP et de Capgemini sont positionnées sur des vRack différents. Il n'y a donc aucune communication possible entre elle.

Pour interconnecter ces 2 environnements, des tunnels VPN IPsec sont créés sur le backbone Internet de OVH entre les firewalls de chacun des ces vRacks (ou réseaux privés virtuels).

Côté ASP, ce VPN est directement porté par les firewalls d'administration C8ASPFIRADM et C9ASPFIRADM.

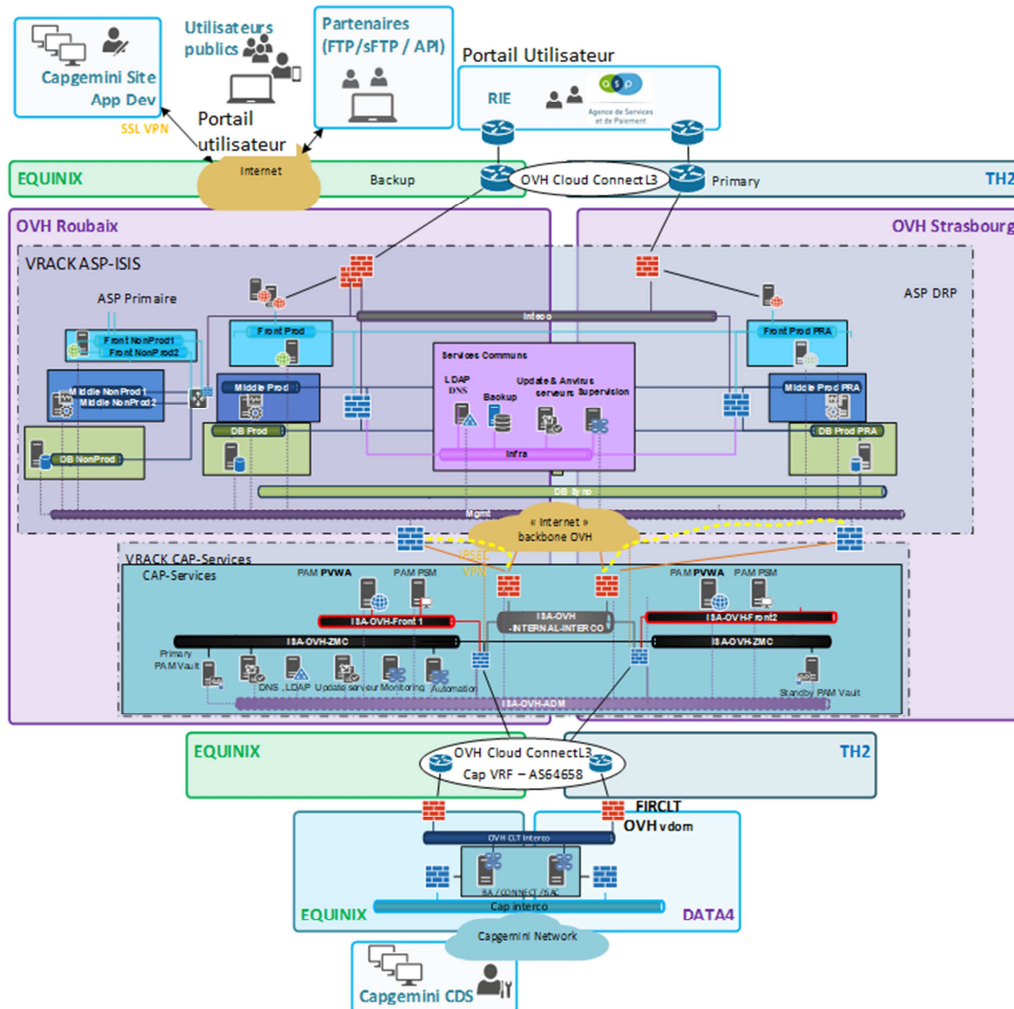
Grâce au service de firewalling d'OVH, les IP publiques de ces passerelles VPN ne sont pas joignables depuis un accès Internet hors OVH.

Ces interconnexions entre les 2 "bulles" OVH CapGemini et OVH ASP permettent aux admins CapGemini de se connecter de façon sécurisée depuis la zone de rebond OVH CapGemini, elle-même accessible uniquement depuis la plateforme d'administration CapGemini mutualisée (MATT). Elles permettent également la remontée des informations de supervision Zabbix, la mise à disposition de dépôts utilisés pour les mises à jour (Redhat Satellite, WSUS, Nexus) et du service de signatures AntiVirus DeepSecurity.



3.2 Schéma de l'architecture

3.3 Listes des environnements



3.3.1 Machines physiques du stockage à froid

1 serveur baremetal OVH est installé sur chacun des sites de Roubaix et Strasbourg. Le serveur de Strasbourg est le serveur nominal.

Ces serveurs sont associés au stockage froid du DataLake.

Ils possèdent une carte RAID matérielle et des disques locaux de type SAS, configurés en RAID5. sur un OS RedHat 7.

Le stockage est partagé en NFS pour les VM DataLake.

CPU : 2x Intel Xeon 4214R - 12 cores
RAM : 96 Go

Les 2 serveurs sont dans la liste des machines en Annexes.

Bien que non implémenté, à date, la fonctionnalité de stockage à froid couvre le use case des grappes de données.

3.3.2 Machines virtuelles

Les VM sont créées par clonage d'un template RedHat 7 minimal, puis elles sont mises à jour via les dépôts du Satellite Redhat. Un ensemble de playbooks de post installation est appliqué pour finaliser leur configuration (réseau/FS/sécurité/comptes/DNS/NTP/etc.)

La liste des VM est disponible en annexes. En plus de cette liste il faut ajouter les réplicas ISIS suivant :

- Réplica Hors-Production utilisé par l'environnement Dev/Rec/Int
- Réplica Production utilisé par l'environnement Pré-production et Production (un pour chaque environnement)

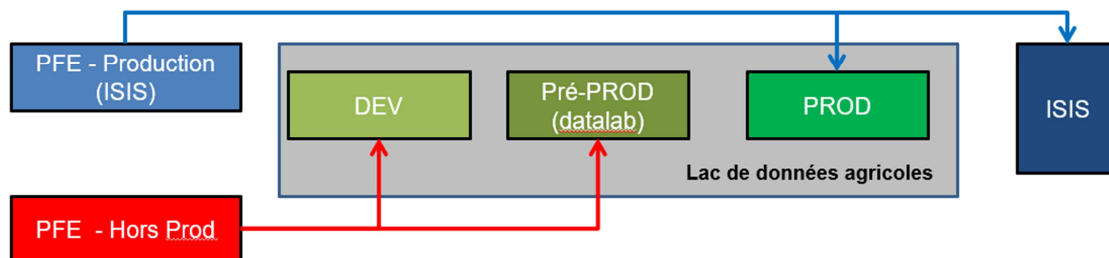
3.3.3 Points de montage

En annexe est disponible la liste des points de montage concernant les besoins spécifiques du Datalake. Les tailles sont exprimées en Go.

3.3.4 Plateforme d'échanges (PFE) de fichiers

Afin de pouvoir échanger des fichiers entre les plateformes hébergées chez OVH avec des partenaires extérieurs à ces plateformes via le RIE (à privilégier) ou via Internet, 2 serveurs virtuels SFTP sont en place : 1 pour la production (mutualisé avec ISIS) et 1 pour les environnements hors production.

- PFE Production reliée avec l'environnement de Production uniquement ;
- PFE Hors Prod pour les phases de recette ou de tests de performance + pour le Data lab.



Les partenaires fournissent leur clé publique SSH et leur(s) IP publique(s) de sortie afin de pouvoir filtrer/sécuriser les connexions (entrantes uniquement) à la PFE.

Un volume de 2 Go est alloué à chaque partenaire (1 prod et 1 horsprod) sur les serveurs physiques de stockage à froid Datalake.

Ces volumes sont partagés en NFS et montés sur les serveurs SFTP et sur les serveurs d'applications Datalake qui servent à produire ou consommer les fichiers échangés.

Sur les serveurs SFTP, ces volumes montés en NFS servent de "home directory" à l'utilisateur dédié du partenaire (un seul compte/login par partenaire, avec si possible une clé SSH différente entre prod et horsprod). Ces "home directory" sont cloisonnés ("jail" / "chrooté"), c'est à dire que le partenaire ne peut naviguer que dans son arborescence et n'a pas de visibilité sur les autres arborescences.

- Des sous répertoires par environnements cible sont créés dans chaque volume :
 - "prd" en production
 - "ppd" et "rec" en hors production
- Liste des partenaires (juillet 2023) * :
-

- DDRP
- FAM
- ODARC
- DSDA

* la liste des partenaires est à titre d'exemple et non exhaustive. Se référer aux spécifications projet pour connaître les partenaires pour lesquels la PFE est utilisée.

- Exemple du volume pour le partenaire ODARC monté sur un serveur SFTP (utilisateur "synapse_odarc") :

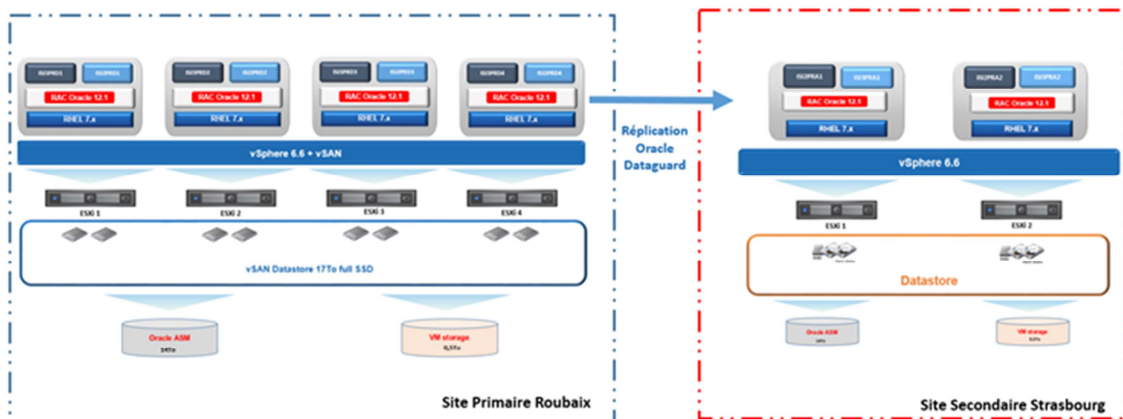
Filesystem	Used	Avail	Use%	Mounted on	Size
zzz.12:/exports/DTL_prod_pfe_odarc	2.0G	6.0M	1		
.8G 1% /jail-odarc/home/synapse_odarc/SYNAPSE_ODARC					

3.4 Bases de données

Se référer au paragraphe §4.7 'Bases de données' du document [DA][10].

Les bases de données de production sont installées en cluster RAC, plusieurs nœuds de ressources CPU participent au cluster dans la limite des licences acquises pour ISIS. Une réplication est faite de façon automatique via DataGuard (en mode asynchrone) sur le site secondaire. Si la latence inter-sites le permet par la suite, et sans impacter les performances des bases de production, le mode de réplication pourrait être passé à synchrone.

Les sauvegardes sont gérées par RMAN, en local à Roubaix, et répliquées à Strasbourg.



Les données sont donc présentes sur le 2^{ème} site, en cas de PRA.

- 2 nœuds en RAC seront utilisés pour l'environnement de pré production PRC.
- Les bases des autres environnements seront hébergées sur 1 seul serveur, distinct par environnement.

3.5 Composants et services

3.5.1 Load-Balancing

Le partage de charge du trafic en provenance d'Internet vers les frontaux Web se fait grâce à la fonctionnalité de Load-Balancing native aux Fortiweb.

3.5.2 Configuration homogène - Ansible

Pour information, les machines virtuelles sont provisionnées de façon automatisée et industrielle grâce à Ansible.

Toutes les configurations post-installation sont également effectuées via Ansible, dans le cadre de l'administration systèmes.

Les journaux d'actions (« playbooks ») Ansible permettent de rejouer les mêmes gestes et d'avoir un résultat cohérent et homogène sur toute la plateforme.

Le serveur principal Ansible est situé dans le réseau Capgemini et pilote de façon sécurisée avec les interfaces d'administration des serveurs de la plateforme ISIS.

Un accès SSH via une clé dédiée est utilisée par le compte 'Ansible' avec des droits root via « sudo » sur tous les serveurs de la plateforme ISIS. Il n'y a pas d'agent à déployer.

Même si aucun engagement d'automatisation du déploiement de l'infrastructure n'est requis, l'objectif est de simplifier les déploiements et l'exploitabilité de la plateforme.

Cette automatisation de déploiement de l'infrastructure est décorrélée la partie automatisation des déploiements logiciels et mises en production. Elles n'ont pas les mêmes contraintes contractuelles.

3.5.3 Sauvegarde et réplication VEEAM

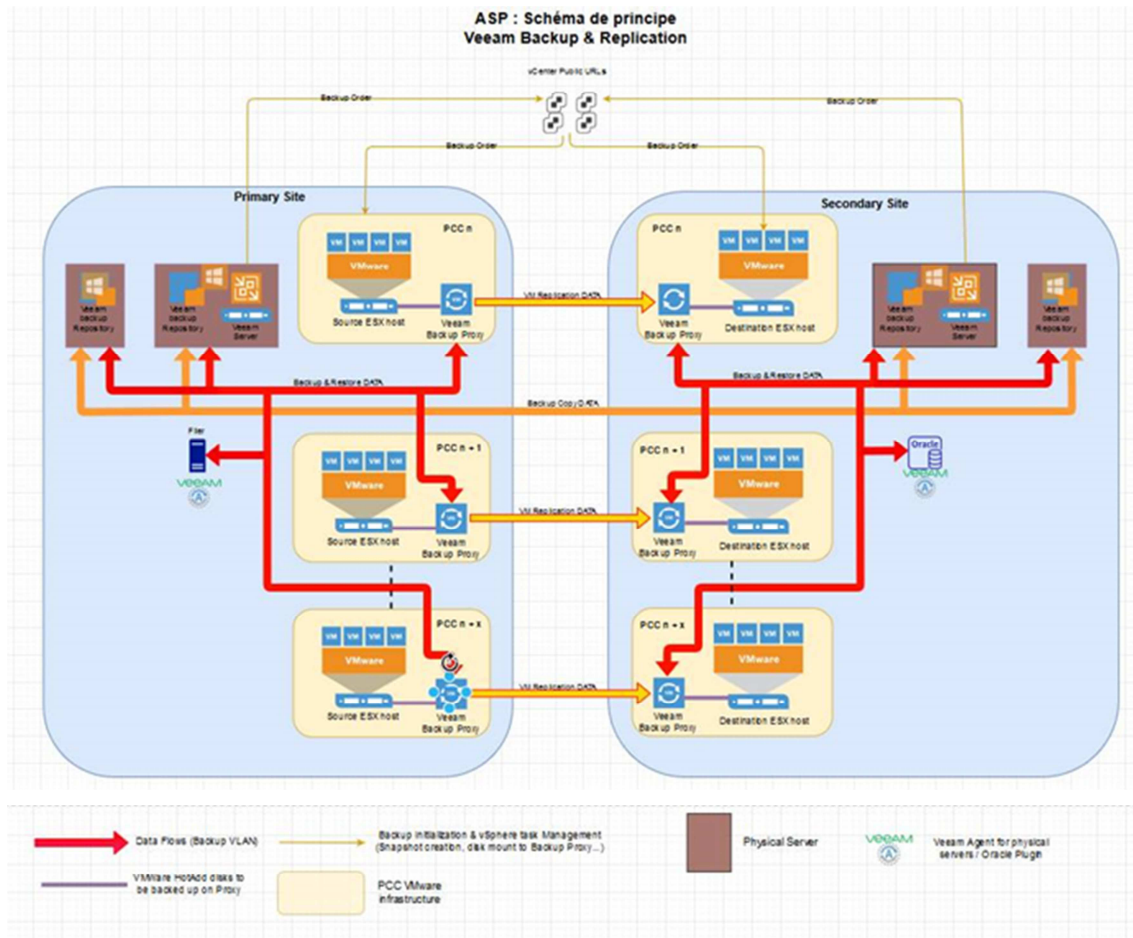
Un des éditeurs majeurs d'outils de sauvegarde. Leur suite logicielle éponyme est utilisée pour la sauvegarde de tout le périmètre. **[DA][12]**

Un serveur maître Veeam est installé sur les 2 sites, Roubaix et Strasbourg.

Les sauvegardes principales sont faites sur Roubaix (site principal) puis répliquées sur Strasbourg (utilisé en cas de PRA), le tout étant piloté par Veeam.

Un proxy virtuel est installé sur chaque cluster de hôtes virtuels, il permet de sauvegarder les machines virtuelles sans avoir besoin d'installer d'agents sur celles-ci.

Les dépôts (ou « repositories ») stockent les sauvegardes. Ce seront des machines physiques afin de pouvoir fournir les espaces disque suffisants et les externaliser des clusters de virtualisation. La réplication de ces dépôts entre les 2 sites est également gérée via VEEAM.



3.5.4 Concentration des logs système - GrayLog

Composant open-source qui permet de collecter/centraliser les logs systèmes (via syslog).

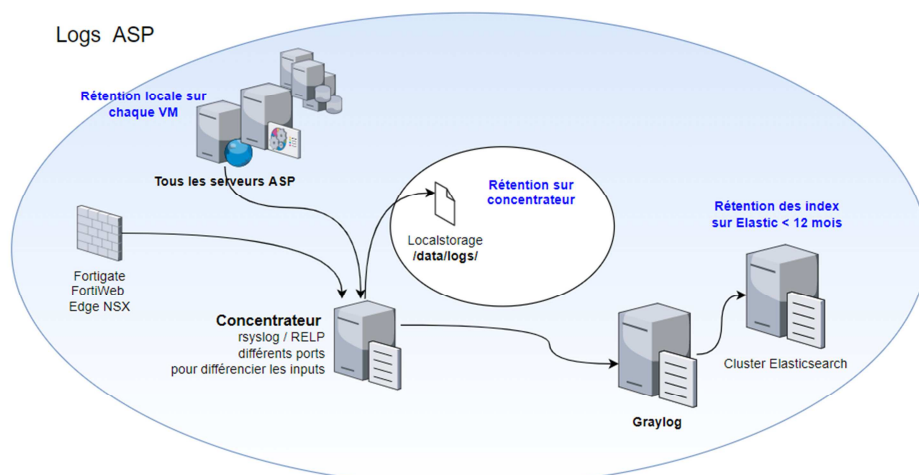
Tous les serveurs et équipements de la plateforme envoient leurs logs système sur ce composant, de façon sécurisée.

Il est placé dans la zone commune de l'infrastructure ISIS.

Des inputs/streams spécifiques sont créées pour les logs applicatifs Datalake.

Les VM graylog sont sauvegardées comme les autres VM.

Il existe 2 puits de logs : production et hors production



3.5.5 Supervision Zabbix

Les équipements des environnements du périmètre sont supervisés par Zabbix. L'ensemble est décrit dans un document spécifique **[DA][11]**

Un client zabbix est installé sur chaque serveur de l'infrastructure. Il envoie ses collectes et alarmes de façon chiffrée à un proxy zabbix installé dans le VLAN « commun » de la zone client via leur interface d'administration. Le proxy communique ensuite avec un serveur maître situé dans la zone mutualisée afin de remonter les données vers une console de pilotage.

Les équipements qui ne permettent pas l'installation d'un agent (comme les appliances de sécurité notamment) sont supervisés par SNMP (ou API).

3.5.6 Antivirus / Antimalware

La solution Trend Micro DeepSecurity (version 12) est déployée sur la plateforme.

Cette solution permet de surveiller les activités des machines virtuelles.

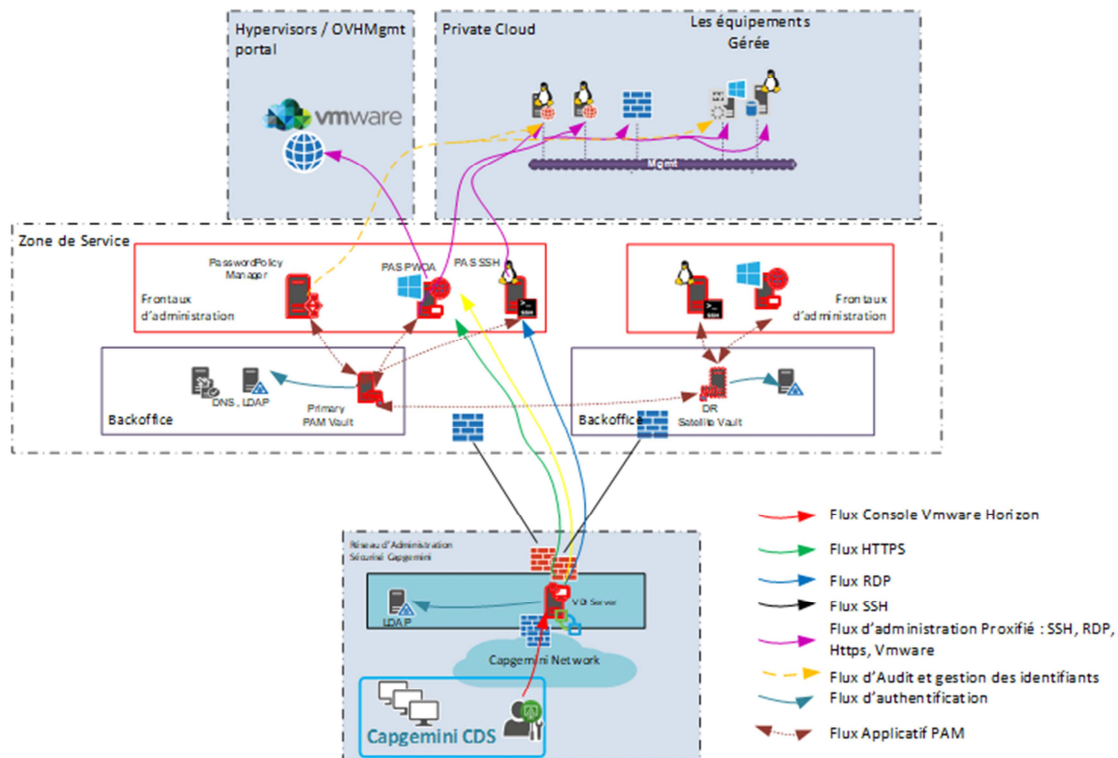
Un client est installé sur chaque serveur pour permettre l'utilisation de la fonction antimalware du produit sous Linux (non disponible via leur solution appliance sur ESXi). Un agent avec des fonctions supplémentaires avancées comme la détection d'intrusion est installée sur les machines « frontales » (web / apache / sftp) situées dans les VLAN accessibles depuis Internet (toujours à travers un firewall).

Les alertes sont remontées sur une console centralisée sur le réseau d'administration Capgemini.

3.5.7 Bastion - CyberARK

Solution logicielle de bastion, qui contrôle, trace et gère les accès aux infrastructures. Ce bastion est à usage Capgemini et mutualisé avec tous les clients utilisant les services OVH.

La documentation d'implémentation du bastion ne peut être communiquée sans précautions particulières. Elle détaille des éléments de sécurité de la zone de services mutualisés Capgemini sur OVH.



3.5.8 LDAP - ForgeRock Directory Server

Annuaire LDAP, version 6.5.

Utilisé pour l'authentification des utilisateurs et des services sur les serveurs de l'infrastructure ISIS et Datalake.

A noter qu'il faudra statuer sur le principe :

- Réplication des annuaires entre les différents environnements (PROD, PRE-PROD et MOE) => solution envisagée mais non activée au 30/11/2022.
- Autonomie des annuaires (pas de synchronisation des environnements PROD, PRE-PROD et MOE)
- Réplication des annuaires sur des entités LDAP spécifiques (Groupe) entre les environnements.
Ce point est à statuer d'ici début janvier 2023. Par défaut, c'est l'option 2 qui est mise en place.

Actuellement l'option 2 est appliquée, à savoir:

- Autonomie des annuaires (pas de synchronisation des environnements PROD, PRE-PROD et MOE)

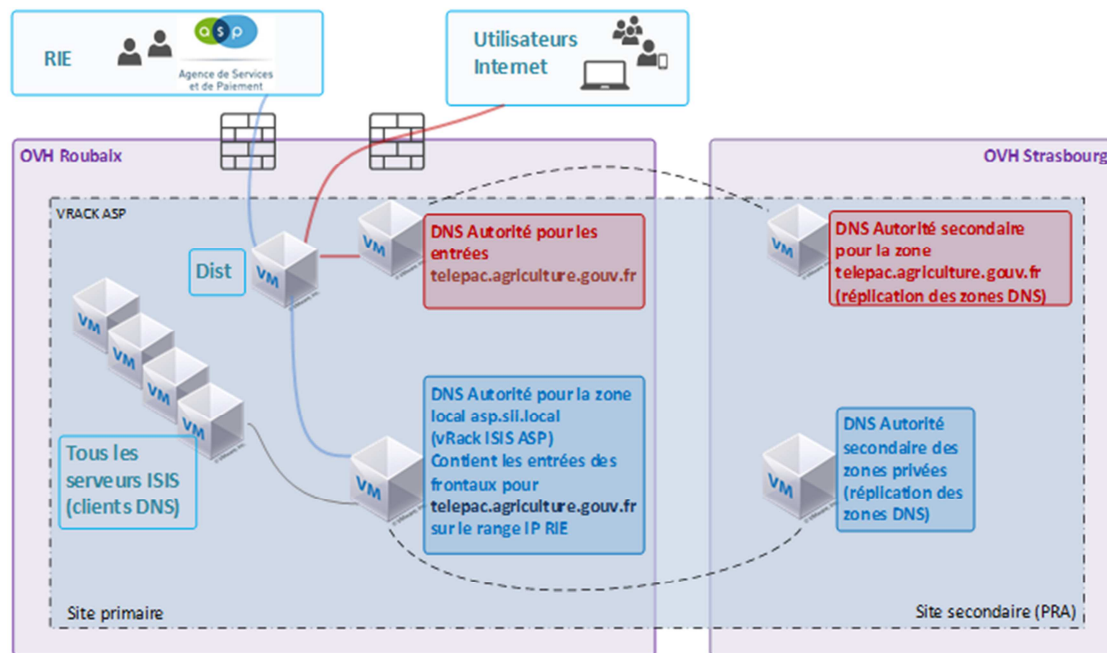
La synchronisation du LDAP système et du LDAP applicatif est réalisée 3 fois par jour.

3.5.9 Service NTP

3 serveurs de temps sont disponibles sur la plateforme (2 sur Roubaix, 1 sur Strasbourg). Chaque serveur de la plateforme est synchronisé avec ces serveurs. Les 3 serveurs de référence sont synchronisés sur des serveurs NTP publiques du groupe « fr.pool.ntp.org ».

3.5.10 DNS

La plateforme possède plusieurs services DNS basés sur la solution PowerDNS (version 4).



Les fonctions authoritative, recursor, ainsi que dnssdist sont utilisées pour rediriger les requêtes de façon appropriée entre le réseau local, RIE et Internet. Les services sont tous dissociés pour simplifier l'exploitation (1 serveur par fonction, fonction sur le port 53 par défaut).

3.5.11 Gestion des mises à jour

3.5.11.1 Mises à jour matérielles

Les mises à jour matérielles (firmware notamment) sont gérées de façon proactive par OVH.

OVH nous avertit lorsqu'un élément doit être mis à jour et nous planifions ensuite ensemble l'implémentation de ces mises à jour.

3.5.11.2 Mises à jour logicielles

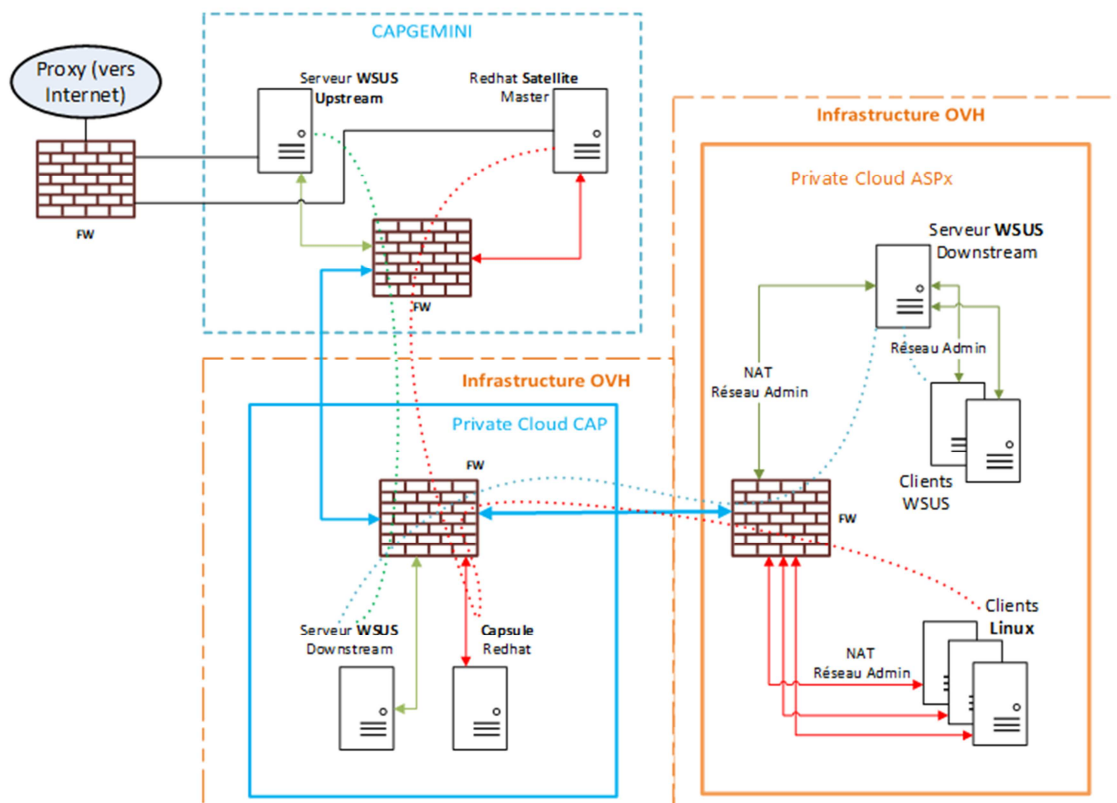
Des campagnes de mises à jour seront menées de façon régulière et au minimum 1 fois par année. Les patches de sécurité seront appliqués tous les 3 mois.

Une plage de maintenance est à définir. Les environnements hors production seront les premiers à être mis à jour, puis la production si aucun impact n'est détecté sur les autres environnements.

En cas de communication ou découverte d'une faille de sécurité critique (score CVSS 8+[\[1\]](#)) une campagne de mise à jour sera effectuée le plus rapidement possible en concertation avec l'ASP.

Les serveurs RedHat sont connectés à un serveur Capsule RedHat Satellite mutualisé (situé dans la zone d'administration OVH-Capgemini) permettant de suivre leur taux de mises à jour, d'évaluer les risques liés aux patches de sécurité et de planifier des mises à jour des serveurs. Cette capsule est reliée au serveur principal RedHat Satellite de la zone d'administration sécurisée Capgemini (sur lequel sont provisionnées les licences RedHat).

Les mises à jour des serveurs Windows (VEEAM et bastion) sont gérées et appliquées grâce à des relais WSUS « downstream » du master installé dans la zone d'administration sécurisée Capgemini.



3.5.12 Relais de mails

La plateforme ISIS/Datalake nécessite un service de mails sortants vers des messageries externes publiques ou vers des utilisateurs ASP.

Un serveur de mail (postfix sur RedHat 7) dédié, et sécurisé comme il se doit (DKIM, SPF et DMARC), est mis à disposition sur le vRack ISIS. Les serveurs de l'application peuvent l'utiliser comme relais de mails par défaut.

Celui-ci ne transmet pas directement les mails vers les destinataires. Ils sont transférés vers un partenaire externe (MailJET) qui assure la remise des mails (gestion de réputation, blacklist, renvoi, etc.) et permet notamment d'obtenir du reporting détaillé sur ces transferts.

3.5.13 Spécifiques à l'application

Un chantier de traitement d'obsolescence des composants middlewares est mis en œuvre en
selon les contraintes projet en accord avec ASP.

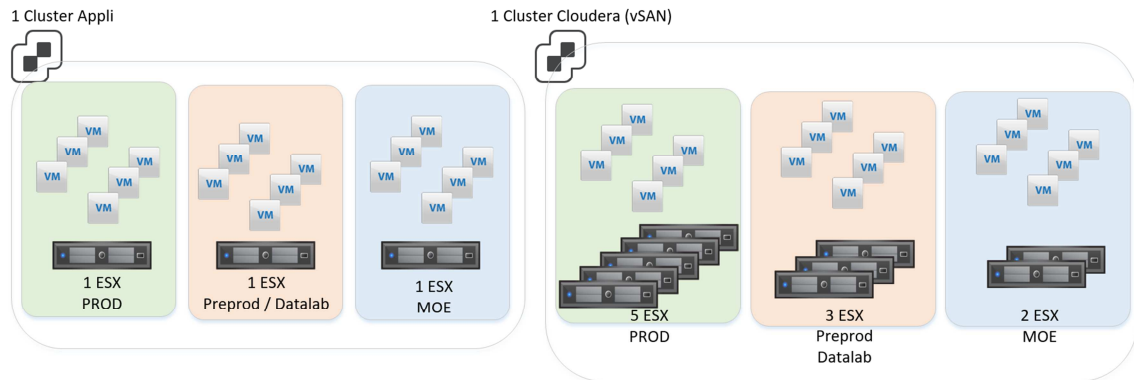
3.5.14 Environnement physique du Datalake

Le Datalake utilise 2 fermes (Private Cloud) spécifiques pour l'ensemble de ces environnements.

- La Ferme Hadoop comprend l'ensemble des machines liées à la plateforme Cloudera.
- La Ferme Applicatif comprend l'ensemble des autres machines pour le Datalake

Des règles d'affinités (DRS) sont appliquées sur ces 2 fermes pour dédier les ESX et isoler les VM des 3 environnements (Prod / PréProd+DataLab / MOE) en terme de RAM et de CPU.

Les VM de chaque environnement sont donc hébergées par des ESX dédiés à 1 environnement. Elles ne peuvent donc plus être réparties sur l'ensemble des ESX.

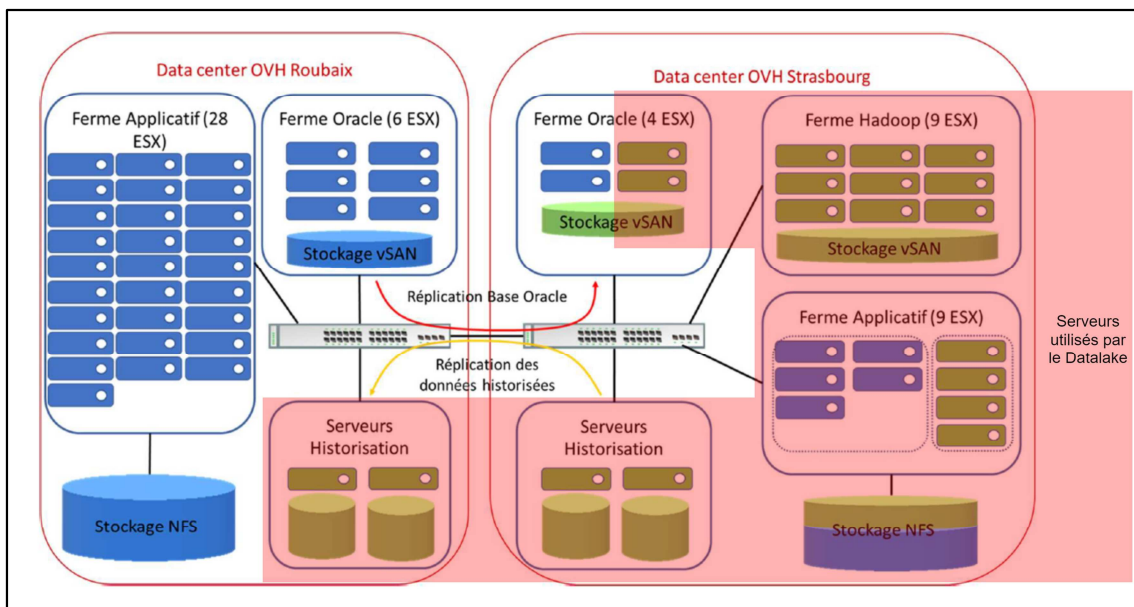


A cela s'ajoute une utilisation de la Ferme Oracle pour les bases répliquées ISIS ainsi que la mise en place de serveur Historisation pour le stockage Froid.

Pour rappel en cas d'incident majeur sur ISIS côté Roubaix les ressources des serveurs du Datalake seront réquisitionnées pour être le PRA ISIS tel que décrit dans la référence [DA9](#).

Les disques utilisés pour le stockage du cluster Cloudera sont des disques NVMe (min 16Gbps) - ils sont locaux sur chaque ESX et sont agrégés/gérés entre les ESX de la ferme par la technologie vSAN de VMware.

Les disques des serveurs de stockage froid sont des disques SAS de 14To (12Gbps) - groupés en grappes de RAID5.

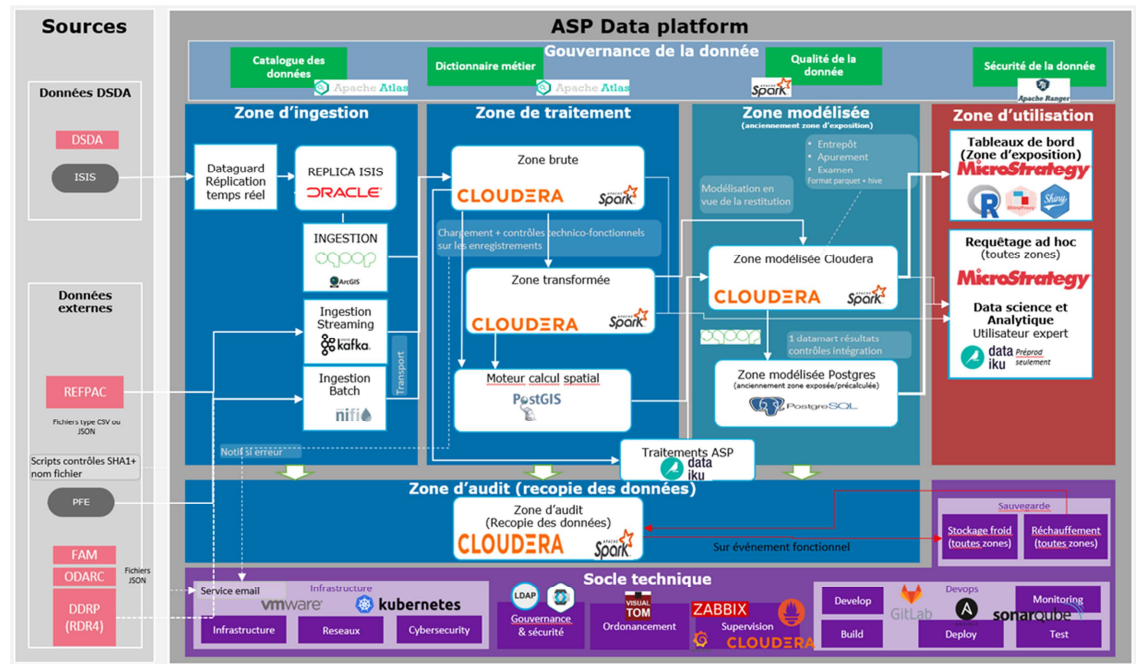


4. Architecture applicative

4.1 Schéma d'architecture applicative

4.1.1 Schéma simplifié

[ASP Data Lake - schéma archi.pptx](#)



4.1.2 Exemple d'articulation entre les composants applicatifs

Via ce paragraphe, nous donnons l'exemple du cycle de vie d'une donnée d'ISIS à destination d'un cas d'usage décisionnel développé sous R-Shiny, ceci pour donner l'articulation entre les différents composants applicatifs du socle (liste non exhaustive).

- Un exploitant effectue une demande d'aide, cette information est donc maintenant présente sous ISIS
- Avec Dataguard, cette information est répliquée en temps réel sur le Réplica ISIS
- Les outils Sqoop, ArcGis, Hadoop Streaming et SqlPlus vont "ingérer" quotidiennement le Replica ISIS pour le déverser dans la partie Données Brutes du Cluster Cloudera.
- Cette donnée brute est maintenant visible de Dataiku pour utilisable pour des cas d'usages Datascience
- La donnée est ensuite transformée puis agrégée avec d'autres données, ceci avec des traitements Spark/Python, pour devenir une donnée modélisée
- Cette donnée modélisée peut ensuite être exposée dans un Datamart PostgreSQL (optionnel : en fonction des cas d'usage)
- Un tableau de bord peut être développé
 - Tableau de bord industriel
 - Sous Microstrategy depuis les données de la couche modélisée Hive(Cloudera)/PotgreSQL
 - Sous Dataiku depuis les données de la couche modélisée pour réalisation de traitements de data préparation industriel, pour

exploitation depuis Microstrategy. cette catégorie de traitement est soumise à validation des équipes Capgemini, afin de s'assurer de l'adéquation avec les principes d'architecture LDA.

- iii. sous R-Studio depuis les données de la couche modélisée PostgreSQL pour produire des rapports dynamique Shiny, soumis à validation des équipes Capgemini
- b. Tableau de bord ad-hoc
- i. Sous Microstrategy depuis les données de l'ensemble des couches LDA: brutes/transférée/modélisée: ce type d'accès n'est pas adapté à du reporting de masse (pb de performance, ...)
 - ii. Sous Dataiku depuis les données de l'ensemble des couches LDA: brutes/transférée/modélisée pour réalisation de traitements de data préparation ad-hoc
 - iii. sous R-Studio depuis les données de la couche modélisée PostgreSQL uniquement pour des analyses statistiques ad-hoc

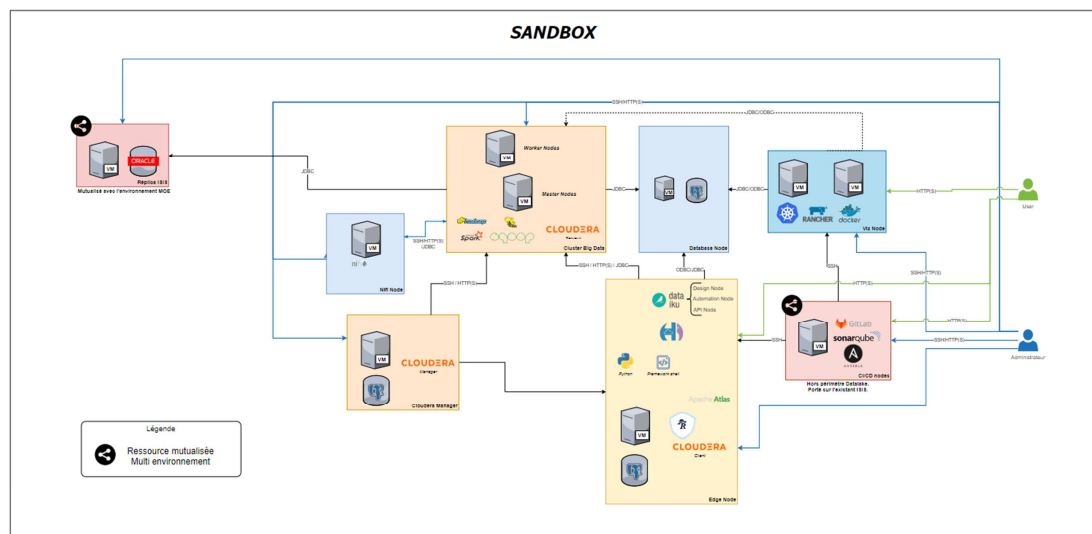
En termes de principes directeur, il est important de respecter le principe de "single source of thruth": les outils de dataviz ne doivent pas intégrer des règles de gestion industrielles.

4.1.3 Schéma d'architecture

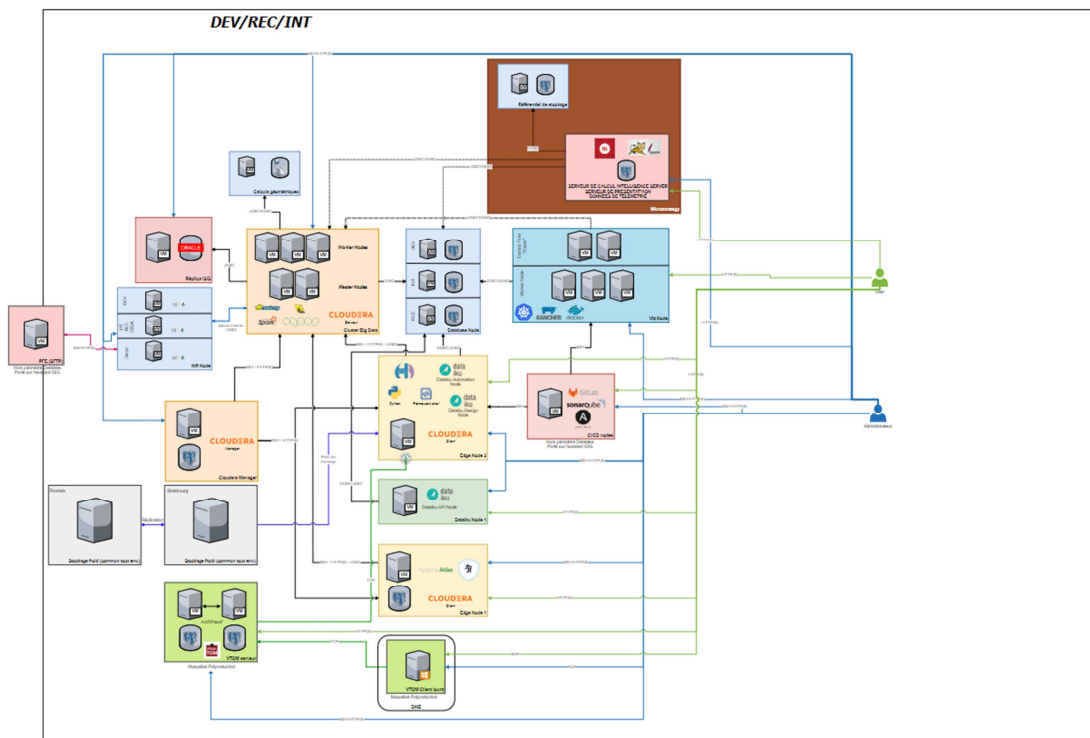
Diagramme Drawio

ASP_architecture_logique_V3.1.drawio

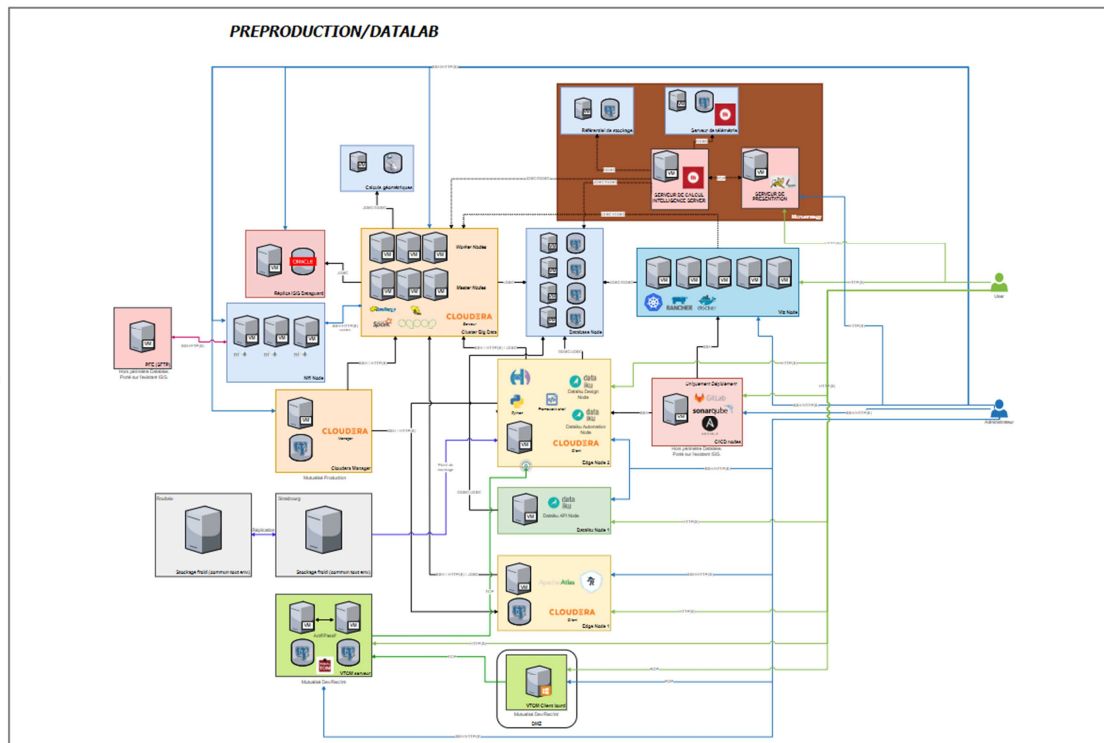
Environnement Bac à sable



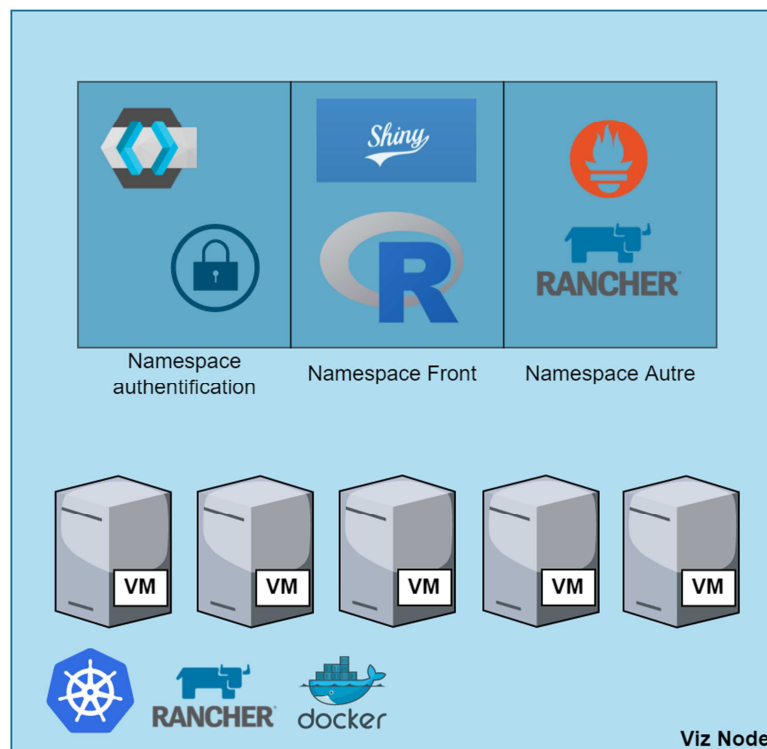
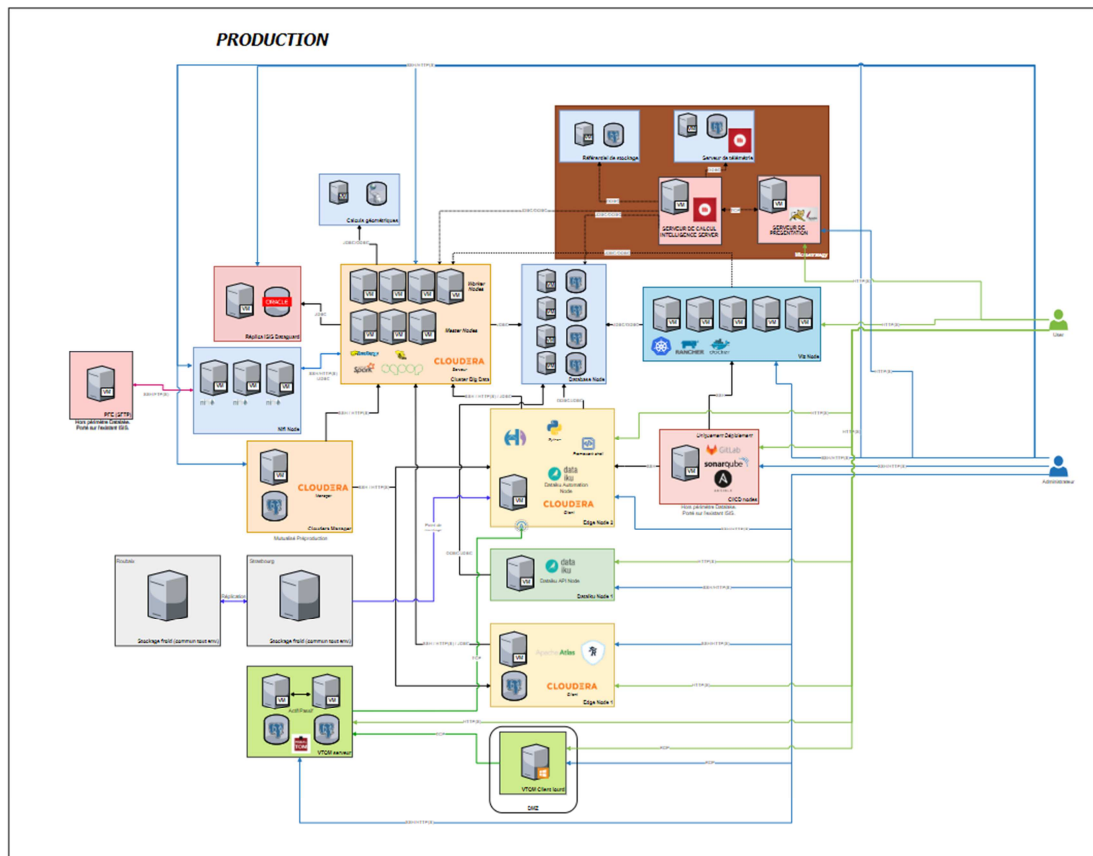
Environnement DEV/REC/INT



Environnement PREPRODUCTION/DATALAB



Environnement PRODUCTION



Zoom Kubernetes

4.2 Contexte du projet

4.2.1 Enjeux métiers

La future réforme de la PAC va confier le soin à chaque Etat membre de fixer précisément les règles et critères d'éligibilité aux aides. A contrario, elle renforcera considérablement le niveau de restitution de données attendue par les autorités européennes. L'apurement des dépenses s'appuiera désormais sur l'atteinte par chaque Etat membre d'objectifs de réalisation et de résultats qui auront été fixés dans le cadre d'un unique plan stratégique proposé par l'Etat membre et validé par la Commission européenne.

Ces objectifs, déclinés sous forme d'indicateurs nécessitent une évolution structurelle du système Isis afin d'améliorer le recueil des données nécessaires à l'élaboration de ces indicateurs, la gestion et la fiabilité des différents indicateurs de réalisation et de résultat.

Enfin, l'objectif global est d'intégrer l'ensemble des données des aides agricoles, incluant les données de la Direction du Développement Rural et de la Pêche (DDRP) de l'ASP.

L'espace DataLab permettrait de favoriser l'innovation en DataScience.

4.3 Exigences fonctionnelles

4.3.1 Exigences fonctionnelles

Les différentes exigences fonctionnelles sont et seront décrites dans les SFD des cas d'usages. A date, la seule spécification disponible concerne le déchargement du Réplica vers la zone brute.

A date, les 3 environnements du Data lake Agricole (MOE/Preproduction et Production) s'appuie uniquement sur la PFE mise en œuvre dans le cadre du marché Accord Cadre ISIS 2020 (mutualisation du service entre ISIS et le Lac de données). L'ensemble des caractéristiques propres à la PFE dans ISIS reste inchangé et sont disponibles dans le DAT ISIS .

4.4 Exigences non fonctionnelles

4.4.1 Contraintes existantes

L'exigence ASP est que les traitements batch (Réplica ISIS et Synapse) n'empiètent pas sur les heures de d'ouverture de services, à savoir 9h – 18h30 (heures métropolitaines).

4.4.2 Volumétrie Réplica ISIS

Quelques éléments d'analyse, à titre indicatif, effectués en Décembre 2021 sur la base Oracle de l'environnement Dev/Rec/Int (ces mesures sont approximatives et le nombre de table oracle et leurs tailles changent selon les campagnes) :

- Nb de tables : 10879
- Taille totale uniquement des données (sans tailles des index) : 4,4 To (pour information les index Oracle sont plus imposants que les données)
- Répartition des tables par taille :
 - Grosse table (≥ 20 Go) → 41 tables pour 2,25 To (51,1% de la volumétrie totale)

- Moyenne table (≥ 5 Go & < 20 Go) → 102 tables pour 0,95 To (21,6% de la volumétrie totale)
- Petite table (< 5 Go) → 10736 tables pour 1,2 To (27,3% de la volumétrie totale)

4.4.3 Objectif de sécurité

4.4.3.1 Authentification

Ces éléments sont décrits dans le document **[DA][6]**

4.4.3.2 Disponibilité

A définir

Critères	Niveaux	Description	Besoins
Disponibilité (D)	D1	Classe de disponibilité = 1 - Pas de remise en cause des services essentiels Interruption de service de 15 jours maximum par an (DI) L'activité peut tolérer une interruption chaque semaine (DO)	
	D2	Classe de disponibilité = 2 - Les conséquences sur les services sont importantes Interruption de service de 3 jours maximum par an (DI) L'activité peut tolérer plus de 10 interruptions par an (DO)	
	D3	Classe de disponibilité = 3 - Les conséquences sur les services sont graves Interruption de service de 8 heures maximum par an (DI) L'activité ne peut pas tolérer plus de 4 interruptions par an (DO)	
	D4	Classe de disponibilité = 4 ou > - Le service doit toujours être fourni Interruption de service de 1 heure maximum par an (DI) L'activité ne peut pas tolérer plus de 1 interruption par an (DO)	

4.4.3.3 Confidentialité

A définir

Critères	Niveaux	Description	Besoins
Confidentialité (C)	C1	Informations pouvant être diffusées ou communiquées à tout public	
	C2	Restriction de la diffusion des informations aux acteurs DGDDI uniquement	
	C3	Informations uniquement accessibles à des utilisateurs identifiés, authentifiés et habilités	
	C4	Informations uniquement accessibles à des utilisateurs identifiés, authentifiés et habilités de manière forte	

4.4.3.4 Intégrité

A définir

Critères	Niveaux	Description	Besoins
Intégrité (I)	I1	Atteinte à l'intégrité du service et des données manipulées acceptée si détectée et signalée	
	I2	Atteinte à l'intégrité du service et des données manipulées acceptée si détectée, signalée et corrigée dans un délai raisonnable.	
	I3	Garantie constante de l'intégrité du service et des données manipulées Atteinte à l'intégrité tolérée si arrêt immédiat des opérations et rétablissement de l'intégrité	
	I4	Garantie constante de l'intégrité du service et des données manipulées Atteinte à l'intégrité non acceptée	

4.4.3.5 Traçabilité

A définir

Critères	Niveaux	Description	Preuve
Preuve (P) Auditabilité	P1	Pas de nécessité d'éléments de preuve	
	P2	Éléments de preuve nécessaires fournis dans un délai raisonnable avec un niveau de traces simple → traces techniques	
	P3	Éléments de preuve nécessaires fournis dans un délai rapide avec un niveau de traces détaillé → traces métier ou fonctionnelles	
	P4	Tous les éléments de trace nécessaire pour garantir la non-répudiation d'une opération	

4.5 Description des composants

4.5.1 Cluster BigData

Le cluster Big Data est basée sur une distribution Cloudera, c'est-à-dire un ensemble cohérent et packagé de composants logiciels ayant pour but la constitution d'une plate-forme Big Data. Il s'agit de la distribution Cloudera Data Platform (CDP) Private Cloud Base de la société Cloudera. Cette distribution sur une suite logiciel Apache et se repose sur un logiciel principal Apache Hadoop.

Les clusters Big Data Hadoop sont définis avec trois type de serveur :

- Master → nœuds maître du cluster, ils contiennent l'ensemble des services nécessaire au bon fonctionnement du cluster Big Data Hadoop
- Worker → nœuds de travail du cluster, ils contiennent les données ainsi que la partie puissance de calcul nécessaire pour les traitements de données.
- Utility-Edge → nœuds de service du cluster, ils contiennent des services annexes du cluster ainsi que les services permettant aux différents outils extérieur de se connecter au cluster Big Data Hadoop

Un noeud supplémentaire de supervision Cloudera est nécessaire pour porter l'outil Cloudera Manager. Celui-ci permet de superviser le cluster Big Data mais aussi gérer les versions logiciels (montée de version / patches / ...).

Les clusters ASP sont définis comme suit :

- DEV/REC/INT → 1 Cloudera Manager/2 Master / 3 Worker / 2 Utility-Edge
- PREPROD/DATALAB → 3 Master / 3 Worker / 2 Utility-Edge
- PROD → 1 Cloudera Manager/3 Master / 4 Worker / 2 Utility-Edge

3 noeuds cloudera manager sont provisionnés:

- 1 cloudera manager dédié au cluster BAC A SABLE
- 1 cloudera manager dédié au cluster MOE
- 1 cloudera manager mutualisé pour les cluster PREPROD/PROD

La technologie Hadoop permet une scalabilité horizontale ce qui signifie qu'en cas de besoin de stockage ou de puissance supplémentaire un nouveau noeud "Worker" peut être installé pour ajouter le nouvel espace ou la nouvelle puissance nécessaire sans impacter l'ensemble de l'architecture déjà en place.

La distribution Cloudera Private Cloud Base 7.1 contient l'ensemble des composants suivants :

Composant
Apache Arrow
<u>Apache Atlas</u>
Apache Calcite
Apache Avro
<u>Apache Hadoop (inclus YARN et HDFS)</u>
<u>Apache HBase*</u>
<u>Apache Hive</u>
Apache Impala
<u>Apache Kafka*</u>
<u>Apache Knox</u>
Apache Kudu
Apache Livy
MapReduce
Apache Ozone
Apache Oozie
<u>Apache ORC</u>
<u>Apache Parquet</u>
Apache Phoenix
<u>Apache Ranger</u>
<u>Apache Solr*</u>

Composant
<u>Apache Spark</u>
<u>Apache Sqoop</u>
<u>Apache Tez</u>
Apache Zeppelin
<u>Apache ZooKeeper</u>

Other Components

Component
Cruise Control
Data Analytics Studio
GCS Connector
<u>HBase Indexer*</u>
<u>Hue</u>
<u>Search*</u>
Schema Registry
Streams Messaging Manager
Streams Replication Manager

Connectors and Encryption Components

Component
<u>HBase connectors</u>
<u>Hive Meta Store (HMS)</u>
<u>Hive on Tez</u>
<u>Hive Warehouse Connector</u>
<u>Spark Atlas Connector*</u>
<u>Spark Schema Registry*</u>

En souligné les services utilisés dans le cadre de la plateforme.

* Composant nécessaire en interne de la suite Cloudera.

L'ensemble de ces composants permet à la distribution de fonctionner correctement. Nous utiliserons principalement les composants liés au stockage (Hadoop, Hive, ...), les composants liés aux traitements de la donnée (Spark, Tez, ...) ainsi que les composants liés à la sécurité et la gouvernance de la données (Ranger, Atlas, ...).

4.5.2 Conteneurisation

Les frontaux applicatifs concernant des usages métiers seront mis à disposition via la mécanique de conteneurisation Kubernetes/Docker via la solution opensource Rancher RKE. Cette mécanique de conteneur est administrée par l'outil Rancher.

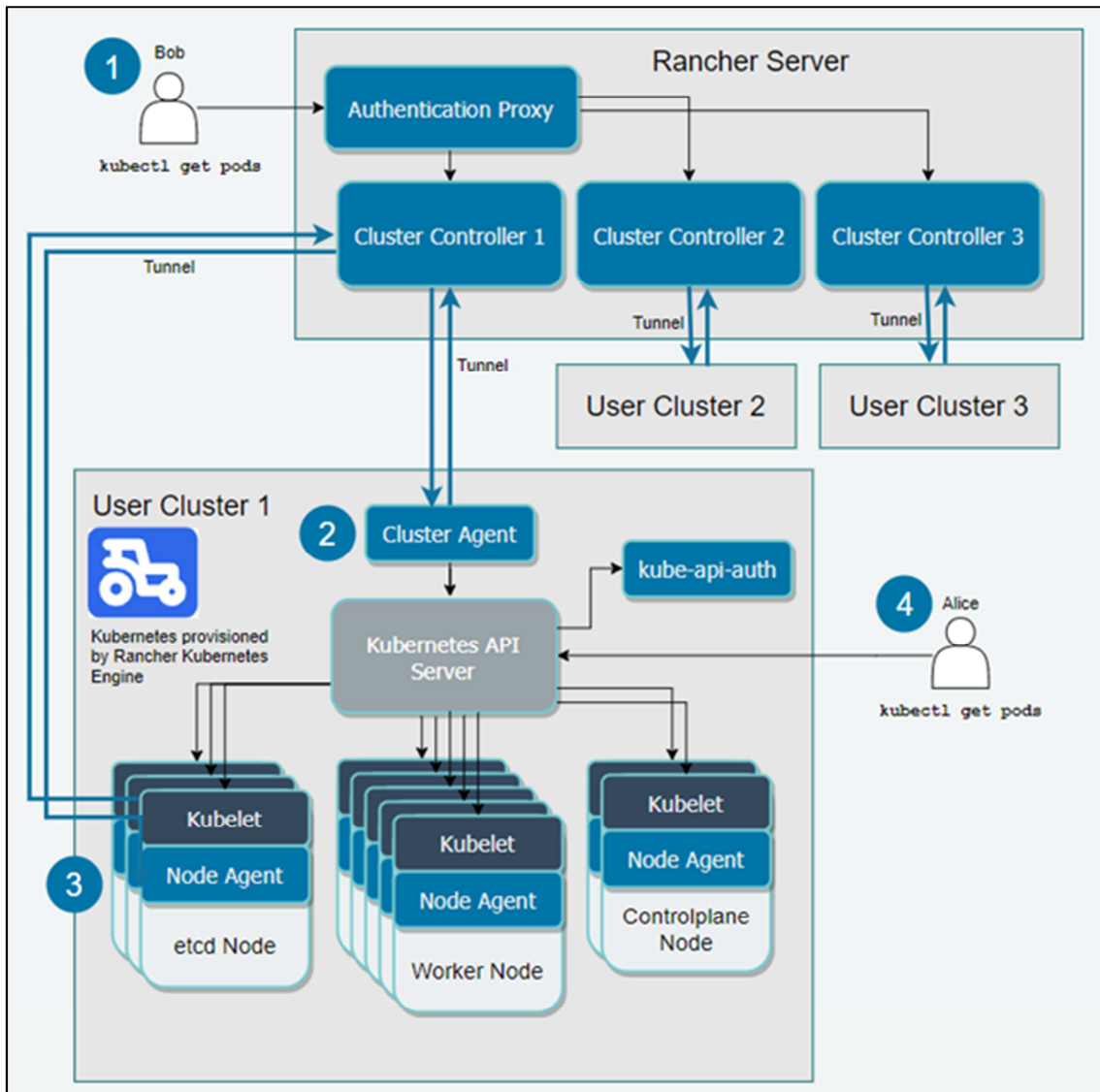


Figure 1 Exemple Cluster K8S et Rancher

4.5.3 Sécurité

CDP utilise Apache Ranger pour fournir une administration et une gestion centralisée de la sécurité au travers d'une interface dédiée (portail Ranger). Les utilisateurs peuvent créer et mettre à jour des règles de sécurité qui sont ensuite stockées dans une base de données sur le cluster. Les plugins Ranger (programmes légers Java) sont intégrés au sein des processus de chaque composant du cluster. Par exemple, le plugin Ranger pour Apache Hive est intégré dans Hiveserver2.

Ces plugins tirent les politiques de sécurité à partir d'un serveur central et les stockent localement dans un fichier. Lorsqu'une requête d'un utilisateur arrive à travers un composant, ces plugins interceptent la demande et l'évaluent au regard de la politique de sécurité. Ils collectent également les données à partir de la requête de l'utilisateur et suivent un thread séparé pour envoyer ces données au serveur d'audit.

Tous les outils utilisant les données sur le cluster BigData passeront par la couche Ranger pour vérifier si l'accès est autorisé.

Les applicatifs de type IHM comme Ranger, Cloudera Manager ou RShiny seront interfacés avec le LDAP/Keycloak comme défini dans le document Authentification. **[DA][6]**

4.5.3.1 Apache Knox

[Apache Knox](#), est mis en place afin de répondre à des besoins de sécurité d'accès aux données du cluster.

L'API Gateway est déployée sur le cluster et permet aux applications externes du cluster d'avoir accès aux différents frameworks d' Hadoop (Hive, Map Reduce, HDFS, etc.) et de les utiliser via un seul canal.

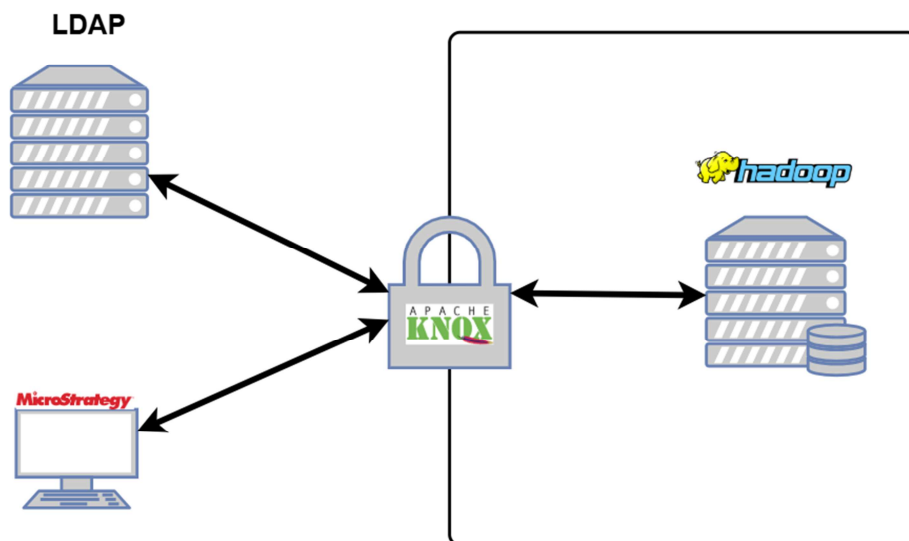
L'initiateur de la requête va dans un premier temps envoyer la requête à Knox qui va interroger le service d'autorisation et/ou d'authentification (tel que Kerberos, Shiro, LDAP, etc.) afin de savoir si cet utilisateur a le droit de contacter le service concerné. Ensuite, Knox va aller interroger un service Hadoop, et retourner la réponse à l'utilisateur.

Cela permet donc de :

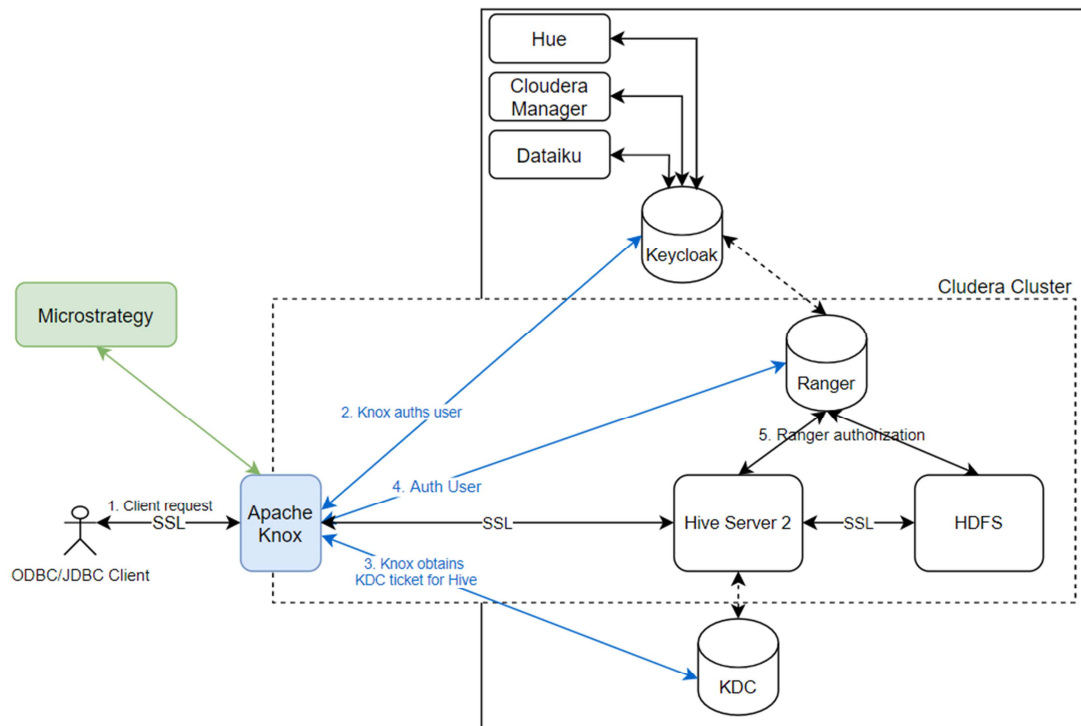
- Protéger les informations du cluster de l'extérieur car on ne s'y connecte plus directement mais via Knox,
- Diminuer le nombre de services avec lesquels le client doit interagir car désormais il communiquera uniquement avec Knox,
- Simplifier le mécanisme d'authentification en n'en utilisant qu'un seul basé sur du HTTP, dans le cas où les clusters sont déjà sécurisés.

C'est d'ailleurs grâce à Knox que l'accès aux données depuis MicroStrategy a pu être établie.

4.5.3.1.1 Schéma simplifié



4.5.3.1.2 Workflow d'authentification Knox



4.5.4 Audit

Apache Atlas est un ensemble évolutif et extensible de services de gouvernance des données et metadonnées.

Les fonctionnalités Apache Atlas sont :

- La classification des données
 - Importation ou définition de la taxonomie orientée métier pour les données
 - Définition, annotation, et automatisation de la capture des relations entre les ensembles de données et les éléments sous-jacents, y compris la source, la cible et les processus de dérivation
 - Exportation des métadonnées à des systèmes tiers
- L'audit centralisé
 - Capture des informations de sécurité d'accès pour chaque application, le processus et l'interaction avec les données
 - Capture des informations opérationnelles pour l'exécution, les étapes et les activités
- La recherche et lignée
 - Chemins de navigation prédéfinis pour explorer la classification des données et des informations d'audit
 - Fonctions de recherche texte pour localiser rapidement et avec précision les données pertinentes et les événements d'audit sur le cluster Datamining
 - Visualisation de l'héritage des données avec des vues drill-down opérationnelles, sur la sécurité et la provenance.
- Un moteur de règles et de sécurité
 - Rationalisation de la politique de conformité lors de l'exécution sur la base de systèmes de classification de données, les attributs et les rôles.

- Définition avancée des politiques pour empêcher la dérivation de données basées sur la classification (par exemple ré-identification) - Interdictions
- Masquage au niveau colonne et ligne basée sur des valeurs et des attributs cellulaires.

4.5.5 Base de données PostgreSQL

La sécurité d'accès aux données dans les BDD PostgreSQL se fera via rôles.

Pour chaque base de données créée 3 rôles sont définis :

- Un rôle Administrateur/Propriétaire avec tous les droits sur la base de données.
- Un rôle Lecteur permettant de lire des données (SELECT).
- Un rôle Utilisateur applicatif pour l'environnement DataLab :
 - Création d'objets (tables, vue, ...)
 - Droits en écriture sur ces objets.

4.5.6 Moteur de calcul spatial PostGis

Une surcouche PostGIS ajoutée à une BDD PostgreSQL dédiée a été mise en place et utilisée comme moteur de calcul géométrique.

L'utilité principale de BDD n'est pas le stockage. Mais la réalisation d'opérations d'analyses spatiales complexes (intersection, différence, mise à niveau, calcul surface,...)

Seules les données nécessaires à ces traitements sont stockées dans cette BDD et peuvent être supprimées en fin de traitement sans que cela n'affecte les traitements du Lac de Données.

Ces données sont principalement de type géométriques "ST_GEOMETRY" associés à des données attributaire (alphanumérique) afin de réduire le nombre d'appels BD pendant le traitement.

4.5.7 Plateforme d'intégration continue (PIC)

Le processus de déploiement est défini dans le document **[DA][3]**.

Nous utilisons les services fournis dans le cadre du projet ISIS. Les outils de PIC utilisés par Datalake sont donc communs avec ISIS (Gitlab, SonarQube, Ansible).

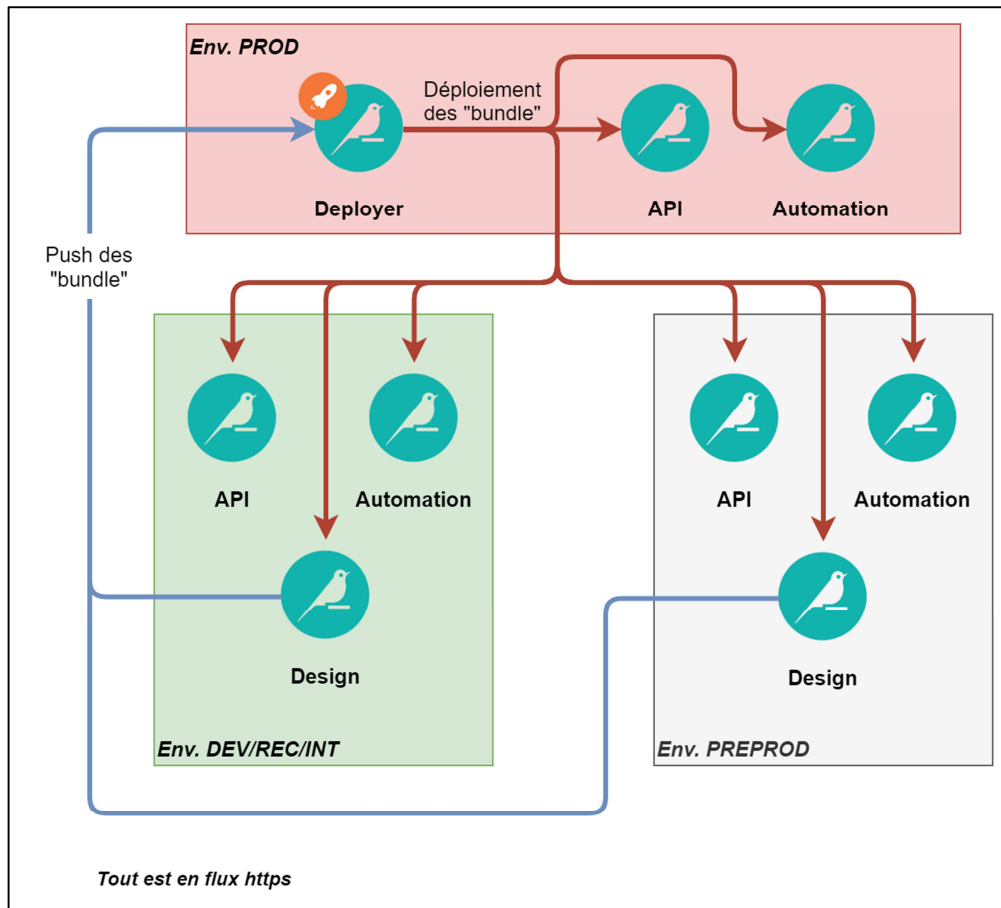
4.5.8 Dataiku

Le logiciel Dataiku est un outil intégrant plusieurs services :

- Design Node → interface de développement
- Automation Node → service d'automatisation des développements, permet d'ordonnancement les traitements développer dans le Design Node.
- API Node → service hébergeant les API développées dans le Design Node
- Deployer Node → service permettant de déployer des "bundles" (package contenant les développements) d'un nœud à l'autre

Les services Dataiku se basant pour certains sur la puissance de calcul du cluster Big Data sont installés sur les nœuds Big Data Hadoop dit "Utility-Edge". Les nœuds "Automation" et "Design" sont ces services. Ils sont tous les deux installés sur le Utility-Edge 2 de chaque environnement.

L'architecture Dataiku mise en place est la suivantes :



Nous avons donc par environnement :

- **DEV/REC/INT**
 - Design (Cloudera Edge 2)
 - Automation (Cloudera Edge 2)
 - API (Dataiku 1)
- **PREPROD/DATALAB**
 - Design (Cloudera Edge 2)
 - Automation (Cloudera Edge 2)
 - API (Dataiku 1)
- **PROD**
 - Automation (Cloudera Edge 2)
 - API (Dataiku 1)
 - Deployer (Dataiku 1)
- La création de connecteur nécessite l'intervention d'un administrateur (Capgemini).
- L'ASP est autonome pour installation de tout package disponible dans Anaconda.
- Toute installation d'une librairie hors Anaconda, fera l'objet d'une validation RSSI Capgemini. Dès lors que celle-ci sera validé, elle sera installée par un administrateur (Capgemini).

4.5.9 RStudio

RStudio permet de couvrir 2 usages:

4.5.9.1 usage industriel de reporting

RStudio permet de publier des applications Shiny pour diffuser des tableaux de bord dynamiques.

Le développement sur RStudio se limite strictement à l'interface utilisateur, sans implémentation directe de règles de calcul.

Les données affichées proviennent exclusivement de la base PostgreSQL déployée dans la zone d'exposition du LDA.

4.5.9.2 usage "datalab"

RStudio peut être utilisé à des fins d'analyse statistique sans besoin d'industrialisation ni déploiement.

Ce mode d'utilisation ne prévoit pas de diffusion/partage des éléments produits depuis RStudio.

Les données manipulées proviennent exclusivement de la base PostgreSQL déployée dans la zone d'exposition du LDA.

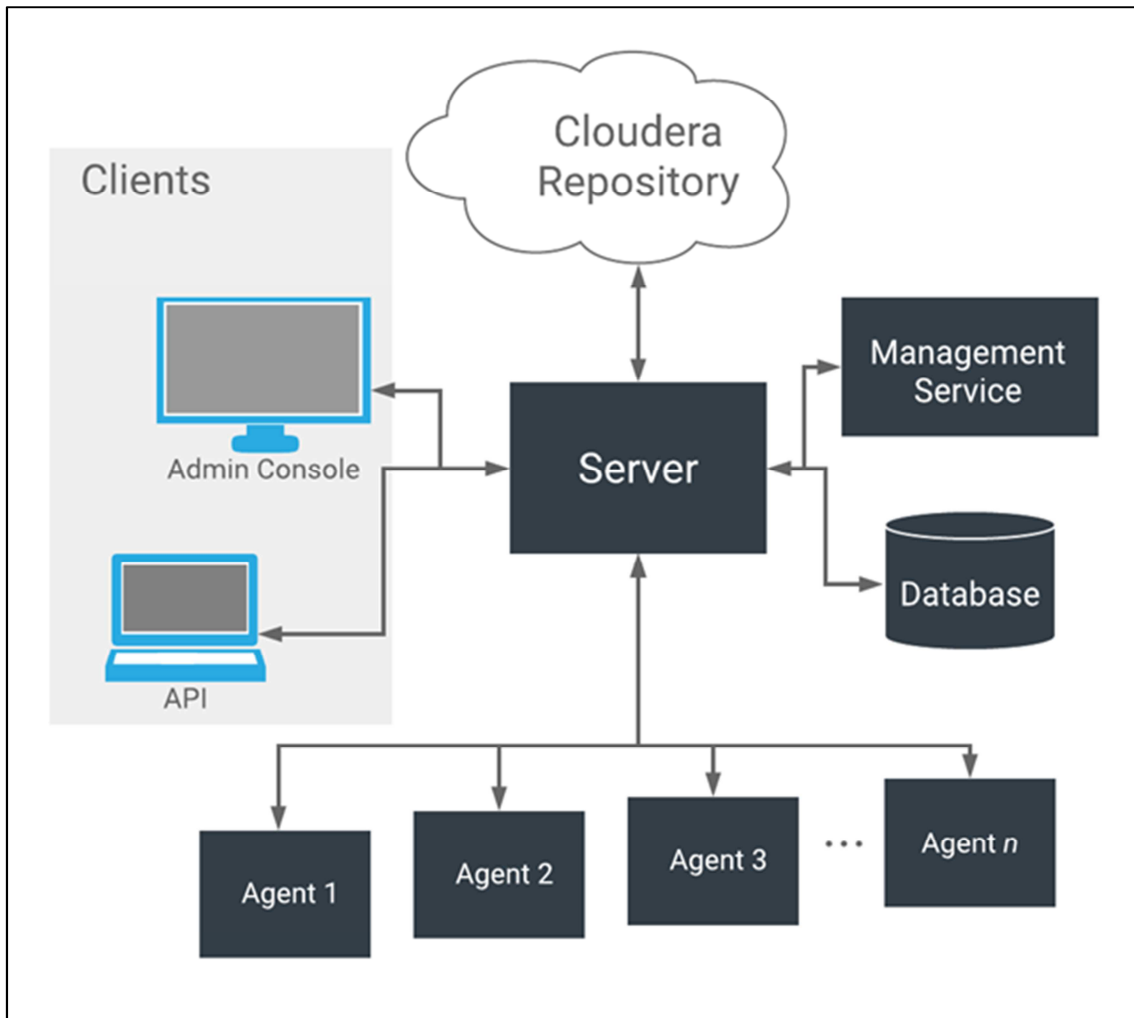
4.5.10 MicroStrategy

le logiciel MicroStrategy est construit autour d'une structure à 3 services:

- Un service de base de données qui abrite :
 - un référentiel de stockage : l'entrepôt de données, qui contient les informations analysées par vos utilisateurs.
 - un référentiel de télémétrie qui contient des informations sur vos projets MicroStrategy.
- Un moteur de calcul **MicroStrategy Intelligence Server**, qui exécute vos rapports, tableaux de bord et documents en se basant sur l'entrepôt de données.
- un **serveur de présentation** construit autour de **Tomcat** qui expose les client légers via les webapps:
 - MicroStrategy Web ,
 - MicroStrategy Library ,
 - MicroStrategy Mobile

4.5.11 Administration

L'administration de la partie Big Data se fait via Cloudera Manager. Cet applicatif permet de manager ainsi que de remonter et d'agréger l'ensemble des services du Big Data. Il centralise l'ensemble des indicateurs fournis par les agents installés sur chaque hôte.



4.5.12 Ordonnancement VTOM

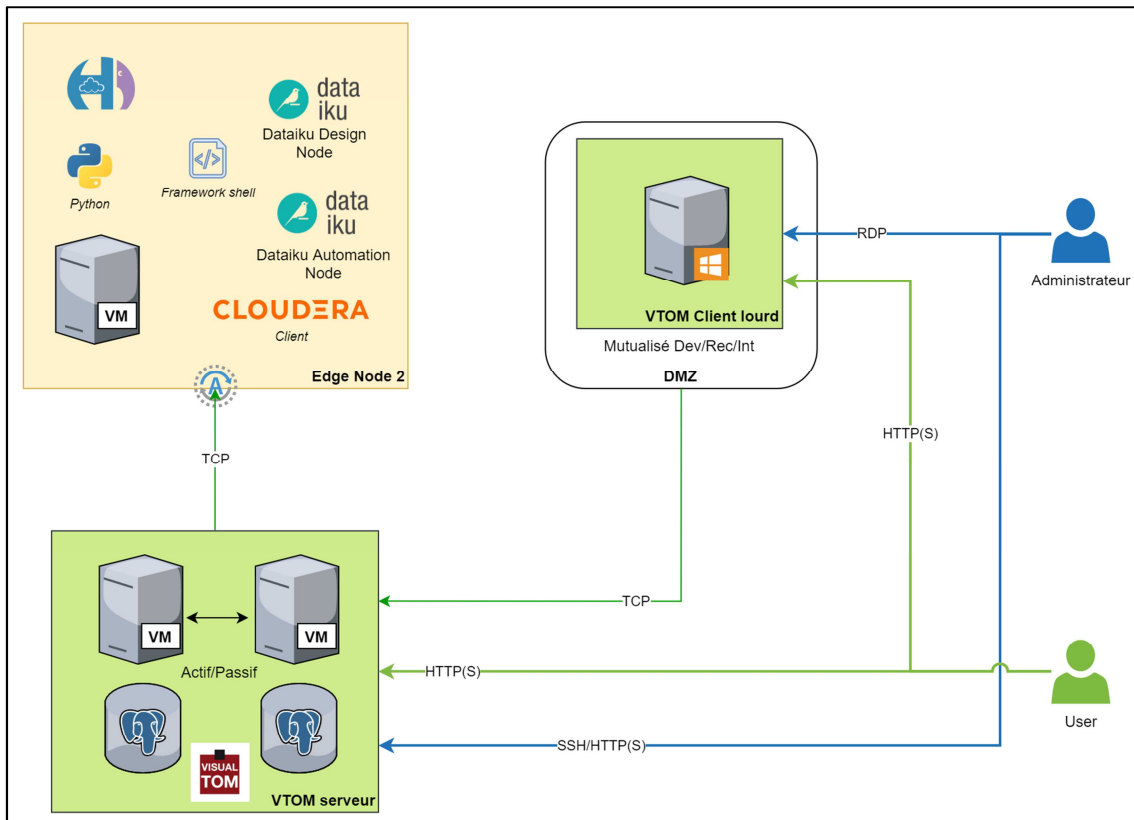
VTOM est un ordonnanceur permettant de cadencer l'ensemble des traitements effectués dans le Datalake. Son architecture est basée sur un système de serveur - agent. C'est-à-dire que les commandes sont pilotées par le serveur et envoyées aux agents pour exécution. Dans le cadre du Datalake, les agents sont installés sur les edge 2 de chaque environnement car les traitements de données sont effectués par ces serveurs.

Deux interfaces sont disponibles, le client léger par protocole HTTPS ou le client lourd par protocole RDP (Remote Desktop Protocol) :

- le client léger permet, selon les droits, de visualiser l'ordonnancement en cours ou passé, les résultats des chaînes de traitements (en erreur ou effectués avec succès) et aussi de monter au plan d'ordonnancement des chaînes de traitement dites "à la demande".
- le client lourd permet, selon les droits, de créer/modifier les chaînes de traitement, de mettre en place un calendrier d'exécution, de monter au plan d'ordonnancement toutes chaînes de traitements

La partie "serveur" de VTOM est en haute-disponibilité par un mode Actif/Passif.

Dans le DataLab, l'ASP pourra planifier des jobs à la demande ainsi que de manière récurrente. Cela sera également possible en recette en concertation avec l'équipe MOE.



4.5.13 Serveurs sur lesquels est installé VTOM

- aspxasvtmp101/aspxasvtmp102 en PRODUCTION
- aspxasvtmh101/aspxasvtmh102 en HORS PRODUCTION

4.5.14 Stockage Froid

Une solution de stockage "Froid" a été mise en place pour stocker les données n'ayant pas nécessité d'être accessibles. Ce stockage est constitué d'une solution de serveur dédié avec un niveau de performance capacitaire adapté et permettant de stocker des données d'archives ou de sauvegarde. Un mécanisme de refroidissement et de réchauffement des données est mis en place dans le cadre du projet. Le stockage "Froid" est vu par le Edge 02 comme un point de montage disque.

A noter : vu en task force le 15/09/2022, il a été indiqué qu'il est possible de lancer un traitement de stockage froid (full) et AUTRE traitement sur le même disque. Par contre un second traitement de stockage froid (full) n'est pas possible.

6. Exploitabilité

6.1 Supervision

Ce chapitre est décrit dans le document **[DA][5]**.

Nous utilisons les services fournis dans le cadre du projet ISIS pour la supervision. L'ensemble des machines sont supervisé par Zabbix et remonté dans l'outillage de ticketing Capgemini ainsi que dans l'outil Grafana ISIS.

6.2 Sauvegardes

Se référer au paragraphe §4.8.4 'Sauvegarde et réplication VEEAM' du document **[DA][10]**.

6.3 Ordonnancement

Ce chapitre est décrit dans le document **[DA][4]**.

6.4 Maintenance

Ce chapitre est décrit dans le document **[DA][1]**.

7. Réponses aux exigences

7.1 Capacités d'accroissement

Se référer au paragraphe §5.1 du document **[DA][10]**.

7.2 Sécurité

Se référer au paragraphe §5.2 du document **[DA][10]**.

7.3 Mise à disposition de logs système

Se référer au paragraphe §5.3 'Mise à disposition de logs système' du document **[DA][10]**.

8. Contraintes de services

8.1 Disponibilité du service

Pour l'environnement de Production

Accessibilité	Plages horaires	Commentaires
Accessibilité depuis l'Internet	7h00 - 19h30 5J/7 hors jours fériés	Sauf intervention planifiée
Accessibilité depuis le réseau RIE	7h00 - 19h30 5J/7 hors jours fériés	Sauf intervention planifiée
Disponibilité de l'infrastructure mutualisée (infrastructures et dispositifs réseaux internes de l'hébergeur)	24h/24 6J/7 hors jours fériés	Sauf intervention planifiée

Pour l'environnement de recette

Accessibilité	Plages horaires	Commentaires
Accessibilité depuis l'Internet	7h00 - 19h30 5J/7 hors jours fériés	Sauf intervention planifiée
Accessibilité depuis le réseau RIE	7h00 - 19h30 5J/7 hors jours fériés	Sauf intervention planifiée
Disponibilité de l'infrastructure mutualisée (infrastructures et dispositifs réseaux internes de l'hébergeur)	24h/24 6J/7 hors jours fériés	Sauf intervention planifiée

9. Annexes

9.1 Liste des VLAN / subnet

9.2 Liste des adresses mises à disposition

3 Liste des VM

Cliquer ici pour afficher le tableau de VMs

Environnement	Description	Hôte	Système d'exploitation	CPU	RAM (Go)
BAC A SABLE					
BAC A SABLE	Cloudera Manager		RHEL 7.9	8	32
BAC A SABLE	Cloudera Master 1		RHEL 7.9	8	32
BAC A SABLE	Cloudera Worker 1		RHEL 7.9	8	32
BAC A SABLE	Cloudera Edge 1		RHEL 7.9	8	32
BAC A SABLE					
BAC A SABLE	PostgreSQL		RHEL 7.9	4	4
BAC A SABLE					
BAC A SABLE	Kubernetes app 1		RHEL 7.9	2	2
BAC A SABLE	Kubernetes app 2		RHEL 7.9	8	12
BAC A SABLE					
BAC A SABLE	Kafka app 1		RHEL 7.9	2	8
DEV/REC/INT					
DEV/REC/INT	Cloudera Manager		CentOS 7.9	8	64
DEV/REC/INT	Cloudera Master 1		CentOS 7.9	16	64
DEV/REC/INT	Cloudera Master 2		CentOS 7.9	16	64
DEV/REC/INT	Cloudera Worker 1		CentOS 7.9	16	128
DEV/REC/INT	Cloudera Worker 2		CentOS 7.9	16	128
DEV/REC/INT	Cloudera Worker 3		CentOS 7.9	16	128
DEV/REC/INT	Cloudera Edge 1		CentOS 7.9	16	64
DEV/REC/INT	Cloudera Edge 2		CentOS 7.9	16	64
DEV/REC/INT					
DEV/REC/INT	PostgreSQL app 1		CentOS 7.9	8	32
DEV/REC/INT	PostgreSQL app 2		CentOS 7.9	8	32
DEV/REC/INT	PostgreSQL app 3		CentOS 7.9	8	32
DEV/REC/INT					
DEV/REC/INT	PostGIS app 1		CentOS 7.9	8	32
DEV/REC/INT					
DEV/REC/INT	Kubernetes app 1		CentOS 7.9	8	32
DEV/REC/INT	Kubernetes app 2		CentOS 7.9	8	32
DEV/REC/INT	Kubernetes app 3		CentOS 7.9	8	32
DEV/REC/INT	Kubernetes app 4		CentOS 7.9	8	32
DEV/REC/INT	Kubernetes app 5		CentOS 7.9	8	32
DEV/REC/INT					
DEV/REC/INT	Kafka app 1		CentOS 7.9	2	8
DEV/REC/INT	Kafka app 2		CentOS 7.9	2	8
DEV/REC/INT	Kafka app 3		CentOS 7.9	2	8

DEV/REC/INT					
DEV/REC/INT	Dataiku 1		CentOS 7.9	8	32
DEV/REC/INT	MicroStrategy				
DEV/REC/INT	MicroStrategy Front & Back End		RedHat 9.4	8	64
DEV/REC/INT	Database Externe		RedHat 9.4	8	64
DEV/REC/INT					
DEV/REC/INT	Docker app 1		CentOS 7.9	8	32
Pré-production					
Pré-production	Cloudera Master 1		CentOS 7.9	16	64
Pré-production	Cloudera Master 2		CentOS 7.9	16	64
Pré-production	Cloudera Master 3		CentOS 7.9	16	64
Pré-production	Cloudera Worker 1		CentOS 7.9	16	128
Pré-production	Cloudera Worker 2		CentOS 7.9	16	128
Pré-production	Cloudera Worker 3		CentOS 7.9	16	128
Pré-production	Cloudera Edge 1		CentOS 7.9	16	128
Pré-production	Cloudera Edge 2		CentOS 7.9	16	128
Pré-production					
Pré-production	PostgreSQL app 1		CentOS 7.9	16	64
Pré-production	PostgreSQL app 2		CentOS 7.9	8	64
Pré-production	PostgreSQL app 3		CentOS 7.9	16	64
Pré-production	PostgreSQL app 4		CentOS 7.9	8	64
Pré-production					
Pré-production	PostGIS app1		CentOS 7.9	8	64
Pré-production					
Pré-production	Kubernetes app 1		CentOS 7.9	8	32
Pré-production	Kubernetes app 2		CentOS 7.9	8	32
Pré-production	Kubernetes app 3		CentOS 7.9	8	32
Pré-production	Kubernetes app 4		CentOS 7.9	8	32
Pré-production	Kubernetes app 5		CentOS 7.9	8	32
Pré-production					
Pré-production	Kafka app 1		CentOS 7.9	6	32
Pré-production	Kafka app 2		CentOS 7.9	6	32
Pré-production	Kafka app 3		CentOS 7.9	6	32
Pré-production					
Pré-production	Dataiku 1		CentOS 7.9	8	64
Pré-production					
Pré-production	Front End		RedHat 9.4	4	16
Pré-production	Back End		RedHat 9.4	8	64
Pré-production	Database Externe 1		RedHat 9.4	4	16
Pré-production	Database Externe 2		RedHat 9.4	4	16
Pré-production / DEV/REC/INT					
Pré-production / DEV/REC/INT	VTOM Serveur 1		CentOS 7.9	2	8
Pré-production / DEV/REC/INT	VTOM Serveur 2		CentOS 7.9	2	8
Pré-production / DEV/REC/INT					
Pré-production / DEV/REC/INT	RDP IHM		Windows Server 2016	16 (cores)	8
Pré-production / DEV/REC/INT					

Pré-production / DEV/REC/INT	Serveur SFTP		CentOS 7	2	4
Production					
Production / Pré-production	Cloudera Manager		CentOS 7.9	8	64
Production	Cloudera Master 1		CentOS 7.9	16	64
Production	Cloudera Master 2		CentOS 7.9	16	64
Production	Cloudera Master 3		CentOS 7.9	16	64
Production	Cloudera Worker 1		CentOS 7.9	16	128
Production	Cloudera Worker 2		CentOS 7.9	16	128
Production	Cloudera Worker 3		CentOS 7.9	16	128
Production	Cloudera Worker 4		CentOS 7.9	16	128
Production	Cloudera Edge 1		CentOS 7.9	16	128
Production	Cloudera Edge 2		CentOS 7.9	16	128
Production					
Production	PostgreSQL app 1		CentOS 7.9	16	64
Production	PostgreSQL app 2		CentOS 7.9	8	64
Production	PostgreSQL app 3		CentOS 7.9	8	64
Production	PostgreSQL app 4		CentOS 7.9	8	64
Production					
Production	PostGIS app 1		CentOS 7.9	8	64
Production					
Production	Kubernetes app 1		CentOS 7.9	8	32
Production	Kubernetes app 2		CentOS 7.9	8	32
Production	Kubernetes app 3		CentOS 7.9	8	32
Production	Kubernetes app 4		CentOS 7.9	8	32
Production	Kubernetes app 5		CentOS 7.9	8	32
Production					
Production	Dataiku 1		CentOS 7.9	8	64
Production					
Production	Kafka app 1		CentOS 7.9	8	32
Production	Kafka app 2		CentOS 7.9	8	32
Production	Kafka app 3		CentOS 7.9	8	32
Production					
Production	VTOM Serveur 1		CentOS 7.9	2	8
Production	VTOM Serveur 2		CentOS 7.9	2	8
Production					
Production	RDP IHM		Windows Server 2016	16 (cores)	8
Production					
Production SBG	Stockage Froid		CentOS 7	2 CPU	96
Réplica Production RBX	Stockage Froid		CentOS 7	2 CPU	96
Production					
Production RBX	Serveur SFTP		CentOS 7	1 CPU	2
Production					
Production	Front End		RedHat 9.4	4CPU	16
Production	Back End		RedHat 9.4	8CPU	64
Production	Database Externe 1		RedHat 9.4	4CPU	16
Production	Database Externe 2		RedHat 9.4	4CPU	16

9.4 Liste des points de montage

Cliquez ici pour développer la liste des points de montage