

DAT

CNRS
BFC & SIRHUS

Document d'architecture Technique du
Socle Technique.

Version 1.9
État : Final

Destinataire(s)

MTES

Historique

Version	Date	Origine de la mise à jour	Rédigée par	Validée par
0.1	01/04/2020	Création	D. MUNIER	
0.2	21/04/2020	Update	D. MUNIER	
0.3	29/04/2020	Update	D. MUNIER	
1.0	11/05/2020	Update	D. MUNIER	
1.1	14/05/2020	Update	D. MUNIER	
1.2	10/06/2020	Update	D. MUNIER	
1.3	15/07/2020	Update	D. MUNIER	
1.4	14/08/2020	Update	D. MUNIER	
1.5	20/08/2020	Update	D. MUNIER	
1.6	06/10/2020	Update	S. NESMOND	
1.7	13/10/2020	Update	D. MUNIER	
1.8	15/03/2023	Mise à jour	Y. MAUGER	
1.9	06/10/2023	Mise à jour migration SNC	Y. MAUGER	

Sommaire

1.	Introduction	5
1.1.	Objectif du document	5
1.2.	Auditoire	5
1.3.	Confidentialité	5
1.4.	Documents applicables	5
1.5.	Définitions et abréviations	5
2.	Solution proposée	7
2.1.	Description du projet	7
2.2.	Description de la solution cible	7
2.3.	Hébergement (Datacenter)	8
2.4.	Plage de Support et Niveau de Service attendu de la plateforme	8
3.	Architecture globale cible	10
3.1.	Hosted Private Cloud SecNumCloud (HPC SNC)	10
3.1.1.	vCenter	11
3.1.2.	Hosts ESX	11
3.1.3.	Tolérance de Panne HPC	11
3.1.4.	Sécurité	11
3.2.	Architecture CaaS	11
3.2.1.	Zones & Environnements	11
3.2.2.	Architecture Applicative	12
3.3.	vRack	12
3.4.	Interconnexions Réseaux	13
3.4.1.	Protection Anti-DDoS	13
3.4.2.	Accès Internet	14
3.4.3.	Zones Réseaux	14
3.4.4.	Routage & Filtrage de la Zone Interne	14
3.4.5.	Load Balancing	15
3.4.6.	NAT	15
3.4.7.	Pare-Feu	15
3.5.	Stockage Objet	16
3.6.	Sauvegarde	16
3.6.1.	Kasten	16
3.6.2.	MIRIA	17
3.7.	DNS	18
3.8.	Service SMTP	18
4.	OpenShift	19
4.1.1.	Présentation	19
4.1.2.	Certificats	19
4.1.3.	Environnements	19

5.	Plan Reprise Activité	23
6.	Supervision	24
6.1.	Prometheus / Grafana	24
6.2.	AlertManager	24
6.3.	Dynatrace	24
6.3.1.	Modes de surveillance	25
6.3.2.	ActiveGate	25
6.3.3.	Principe de fonctionnement de la supervision Dynatrace	25
7.	Opsgenie : Gestion des Alertes et Incidents	27
8.	Fin de document	30

1. Introduction

1.1. Objectif du document

Ce document spécifie l'architecture du projet « MTE - XX'AU » d'un point de vue infrastructure Physique. L'aspect applicatif n'est pas abordé ici (se référer au document annexe suivant : « MTES_XXAU_DAT_V1.8.odt »).

1.2. Auditoire

Ce document est destiné au client « MTE » dans son ensemble, ainsi qu'aux équipes techniques de l'infogérant.

1.3. Confidentialité

Le contenu de ce document est confidentiel (Niveau C3), soumis au secret professionnel ou protégé par la loi. L'utilisation, la copie et la divulgation non autorisées d'une partie ou de l'intégralité de ce document sont susceptibles d'être illégales.

1.4. Documents applicables

Référence	Titre
	Matrice de flux
	Plan d'assurance sécurité
	Document solution
	Plan d'assurance qualité

1.5. Définitions et abréviations

Terme	Définition
Bare Metal	BM - Machine physique dédiée
DDoS	Attaque par déni de service
DMZ	Zone démilitarisée
HA	High Availability (Haute Disponibilité)
HPC	Host Private Cloud d'OVH
IaaS	Infrastructure-as-a-Service
NAT	Network Address Translation
NSX	Réseaux programmables de VMware
PaaS	Platform-as-a-Service
PCA	Plan de Continuité d'Activité
PRA	Plan de Reprise d'Activité
SaaS	Software-as-a-Service
SLA	Service Level Agreement
SOA	Start of Authority (DNS)
TDP	Trusted Digital Platform
VLAN	Virtual LAN
VM	Virtual Machine
vRack	Réseau Privé OVHcloud
vROps	VMware vRealize Operations
VxLAN	Virtual Extensible LAN

HPC	Hosted Private Cloud
RBX	Roubaix
SBG	Strasbourg
GRA	Gravelines
SNC	SecNumCloud

2. Solution proposée

2.1. Description du projet

L'objectif principal du « MTE – XX'AU » consiste en la dématérialisation du processus de traitement de l'instruction des DAU ainsi que l'intégration des formalités complémentaires associées. Il prend en compte les fonctionnalités nécessaires à l'instruction, les services aux usagers et offre des services avals portant sur la totalité des actes et non seulement ceux instruit par les DDT.

Le projet a pour ambition de simplifier les pratiques et de permettre aux agents de privilégier la dimension qualitative à haute valeur ajoutée de leur métier.

2.2. Description de la solution cible

La solution PLAT'AU est une application conteneurisée qui s'appuie sur une plateforme de conteneurs Kubernetes RedHat OpenShift qui permet d'automatiser l'exploitation de toute la pile de gestion des déploiements. L'intérêt de l'utilisation de la containerisation est d'avoir une solution optimisée qui augmente fortement la productivité des développeurs et qui favorise l'innovation tout au long du cycle de vie de l'application.

La plateforme OpenShift est elle-même hébergée au sein d'une solution « HPC SNC » dédiée reposant sur des technologies VMware. Plusieurs tenant OpenShift y sont dédiés aux différents environnement du client MTE.

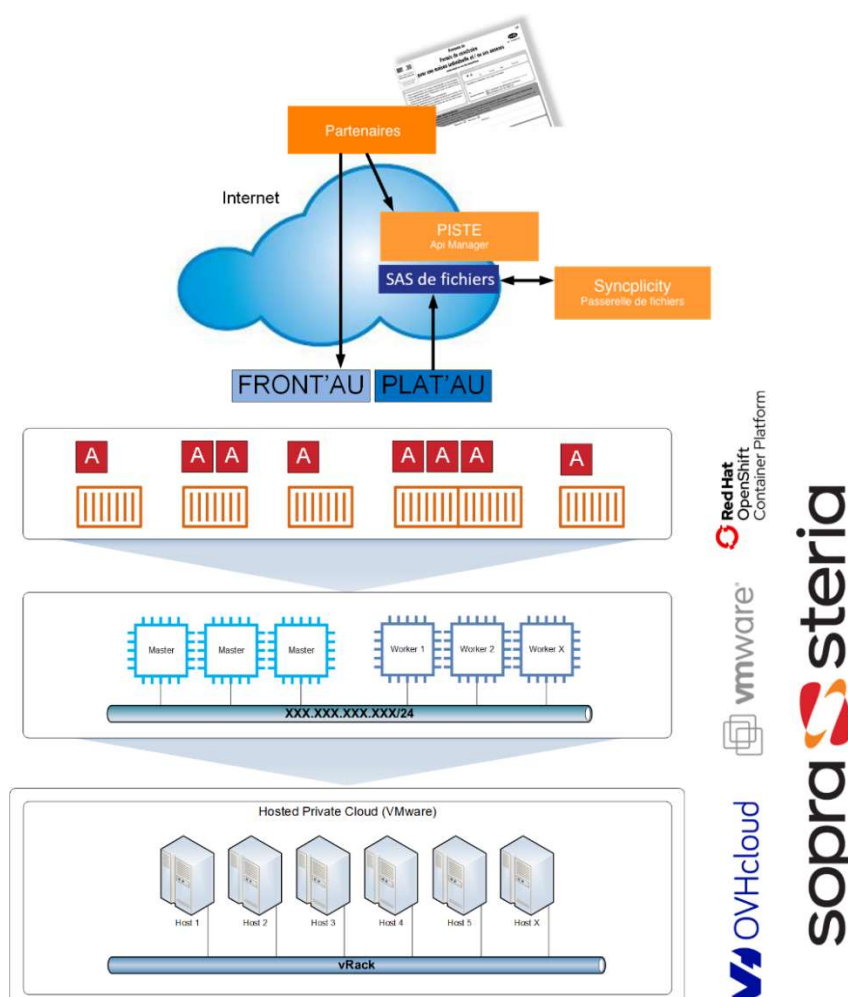


Figure 1 - Schéma de principe de l'application XX'AU : HPC + VMware + OpenShift

2.3. Hébergement (Datacenter)

Le client dispose de deux infrastructures de type « HPC SNC » qui lui sont dédiées et dont l'achat est directement couvert par celui-ci à travers le marché UGAP. Deux régions OVHcloud sont utilisées, respectivement Roubaix (RBX) et Strasbourg (SBG), où se trouvent différents centres de données. Malgré la présence d'OVHcloud également sur le sol américain, la structure juridique de la société n'est pas soumise aux « Patriot ACT » ni au « Cloud ACT ».

Concernant le projet « MTE – XX'AU », l'infrastructure physique est déployée sur les deux régions. Une troisième région, Gravelines (GRA), est notamment utilisée pour externaliser les données de sauvegarde liées à l'outil Kasten.






Figure 2 - Localisation des régions OVHcloud en France

2.4. Plage de Support et Niveau de Service attendu de la plateforme

Pour le projet « MTE – XX'AU », la plage de support et les objectifs en termes de SLA sont les suivants :

- Production :
 - Niveau de Service : Gold
 - Plage de Support : Permanent 7j/7, 24h/24
- Pre-Production :
 - Niveau de Service : Silver
 - Plage de Support : Etendu 6j/7, 6h à 22h
- Hors-Production :
 - Niveau de Service : Silver
 - Plage de Support : Normal 5j/7, 8h à 18h

	 GOLD	 SILVER	 BRONZE
	Disponibilité		
Délai de Remise en Service ¹ (RTO)	2 Heures	4 Heures	8 Heures
SLA Infrastructure	99,8%	99,5%	99%
SLA Applicatif	99,5%	98%	--
	Incident		
Temps d'Acquisition (TTO)	15 Minutes	30 Minutes	60 Minutes
Temps de Résolution (TTS) ²	P1 : 2 Heures P2 : 4 Heures P3 : 8 Heures	P1 : 4 Heures P2 : 8 Heures P3 : 2 Jours	P1 : 8 Heures P2 : 2 Jours P3 : 5 Jours
	Plage de Service Garanti ³		
	Normal : 5 jours sur 7, Hors Sam. & Dim. & jours fériés, de 8:00 à 18:00 Etendu : 6 Jours sur 7, Hors Dim. & jours fériés, de 6:00 à 22:00 Permanent : 24x7		
	GOLD	SILVER	BRONZE
¹ : Remise du service définitif, ou avec un contournement. ² : Temps de résolution ou solution de contournement. ³ : Possibilité de moduler la plage de service avec le niveau de support selon les besoins du client.			

3. Architecture globale cible

3.1. Hosted Private Cloud SecNumCloud (HPC SNC)

La couche physique qui porte la totalité de la plateforme « MTE – XX'AU » repose sur la technologie Host Private Cloud SecNumCloud (HPC SNC) d'OVHcloud. Hosted Private Cloud est une solution PaaS gérée par OVHcloud, qui se base sur la technologie VMware, incluant vSphere, vCenter, NSX ainsi que vROps. HPC combine l'évolutivité du Cloud à une infrastructure matérielle 100% dédiée.

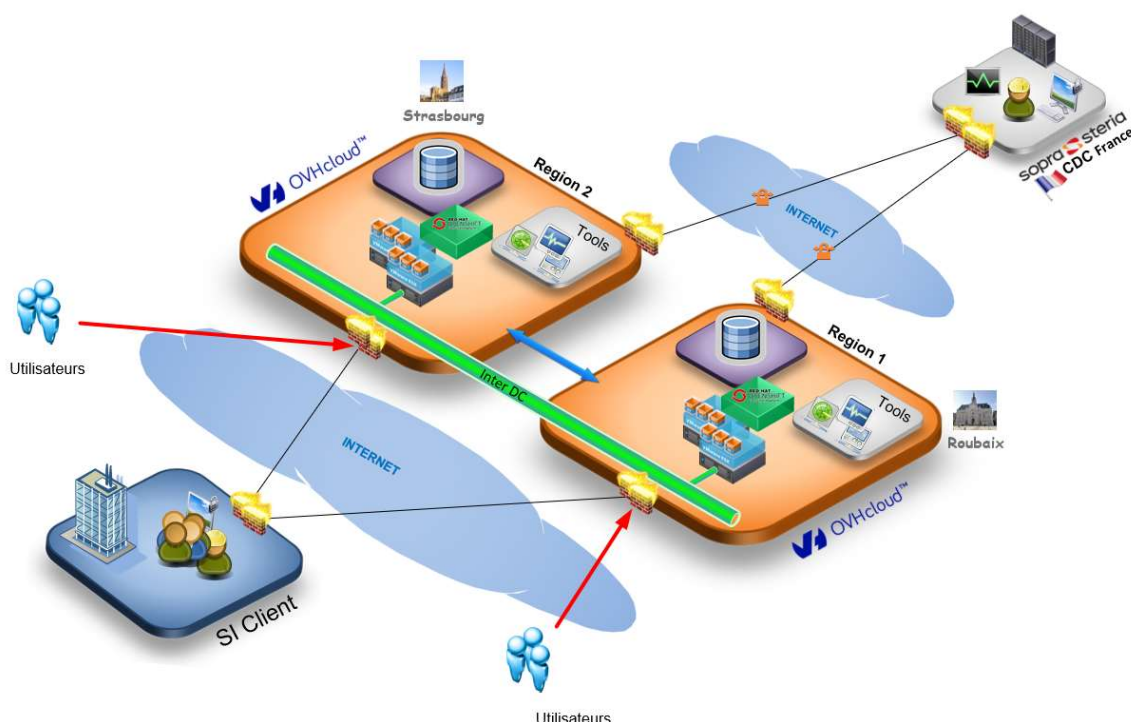
SecNumCloud, référentiel élaboré en 2016 par l'ANSSI, permet la qualification de prestataires de services d'informatique en nuage, l'offre « HPC SNC » d'OVHcloud dispose de cette qualification depuis début 2021.

La solution est hébergée sur deux Cluster « HPC SNC » dédiés au MTE, un premier Cluster vSphere étant déployé à Roubaix et à destination des environnements de Hors-Production, un second étant déployé à Strasbourg et à destination des environnements de Production.

Il est à noter que le site de Roubaix servira de site de PRA, l'ensemble des environnements de Hors-Production qu'il porte en mode de fonctionnement nominal seront arrêtés en cas de déclenchement du PRA. A tout moment, les Cluster « HPC SNC » de Roubaix et Strasbourg devront héberger le même nombre et le même type d'hôtes (ou à défaut des puissances cumulées CPU / RAM équivalentes).

- Environnement « Production » (SBG) : **PCC-51-178-67-78**
- Environnement « Hors-Production » (RBX) : **PCC-51-178-72-98**
- Environnement « Reprise (PRA) » (RBX) : **Voir « Hors-Production »**

A ce jour et pour ce projet, aucun mécanisme de PRA n'est implémenté au niveau de cette couche d'infrastructure HPC, un PRA est néanmoins prévu au projet.



Principe de l'infrastructure technique de la solution

3.1.1. vCenter

VMware vCenter est la plate-forme de gestion des infrastructures informatiques virtuelles fournie par VMware qui permet aux utilisateurs de gérer de manière centralisée et optimisée des environnements virtuels. vCenter offre une gamme complète de fonctionnalités pour le déploiement, l'automatisation, la gestion et la sécurité des infrastructures virtuelles, ainsi que la gestion des performances, la surveillance et la conformité. vCenter apporte également une interface graphique accessible depuis un navigateur Web.

3.1.2. Hosts ESX

Les Hosts ESX sont des serveurs physiques (CPU / RAM / Stockage) avec l'hyperviseur VMware installé « On Top ». Ces ESX sont intégrés dans un HPC SNC qui forme un Cluster afin de supporter les différentes VMs et sont sous la responsabilité opérationnelle de l'infogérant. Le projet « MTE – XX'AU » repose sur l'offre HPC SNC d'OVHcloud en version vCSA 7.0.

S'agissant du projet, deux Clusters HPC sur deux régions OVH différentes sont utilisés, des hosts ESX différents et dédiés à chaque HPC SNC sont déployés sur chacune des régions.

3.1.3. Tolérance de Panne HPC

VMware HA (High Availability) est une fonctionnalité qui vient avec la solution vSphere. Elle permet de redémarrer les machines virtuelles sur un autre hôte du cluster en cas de défaillance matérielle de l'un d'eux. Cette fonctionnalité est activée par défaut.

Le calcul du nombre d'ESX a été effectué de telle manière que la perte d'un ESX puisse être absorbée par les autres. En fonctionnement nominal, le ratio vCPU est fixé à 1:2 sur l'environnement de Production et à 1:3 sur l'environnement de Hors-Production, celui de la RAM étant fixé à 1:1 sur les deux environnements.

3.1.4. Sécurité

Une politique d'accès au vCenter est activée. De ce fait, en plus d'une combinaison « Identifiant + Mot de Passe », l'accès n'est possible que via la mise en place d'une liste blanche d'adresses IP autorisant l'accès à l'interface Web du vCenter ou aux API de gestion du vCenter uniquement à certaines IP.

S'agissant des offres HPC SNC, l'utilisation d'une authentification multi facteur est de plus obligatoire par défaut.

3.2. Architecture CaaS

Le cœur de la solution repose sur la technologie de conteneurisation RedHat OpenShift Platform qui fournit une plate-forme Cloud pour le déploiement, le développement et la gestion de l'application.

La technologie des conteneurs est une alternative à la virtualisation classique, qui permet un gain considérable de ressources (CPU, RAM, Stockage) à l'instar des machines virtuelles qui embarque tout un système d'exploitation.

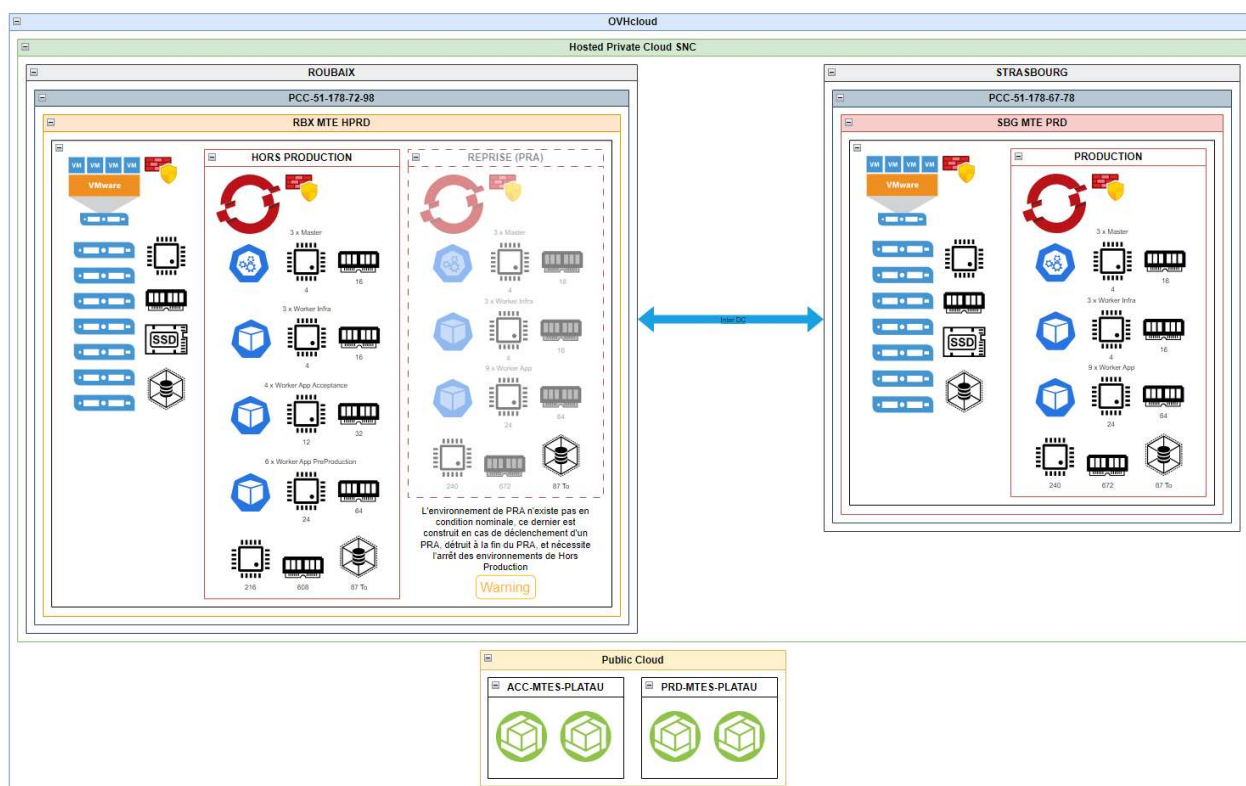
OpenShift est porté par l'hyperviseur VMware. OpenShift propose par ailleurs différents mécanismes d'isolation tels que les espaces de nommage (NameSpace), les Rôles, les Privilèges, les Security Groups... permettant de renforcer la sécurité.

L'intérêt de l'utilisation de la containerisation est d'avoir une solution optimisée qui augmente fortement la productivité des développeurs et qui favorise l'innovation tout au long du cycle de vie de l'application.

3.2.1. Zones & Environnements

Les environnements OpenShift seront répartis suivant le principe décrit en [3.1](#). Sur les deux régions OVH sur lesquelles les Cluster HPC SNC sont déployés l'environnement de Production

sera déployé sur le Cluster HPC SNC de Strasbourg, l'environnement de Hors-Production sera déployé sur le Cluster HPC SNC de Roubaix. Le Cluster HPC SNC de Roubaix hébergera l'environnement de Reprise (PRA) lorsque ce dernier sera déclenché. L'ensemble des environnements de Hors-Production qui y sont habituellement hébergés seront alors obligatoirement arrêtés.



Principe de répartition Production / Hors-Production / Reprise sur les régions OVH

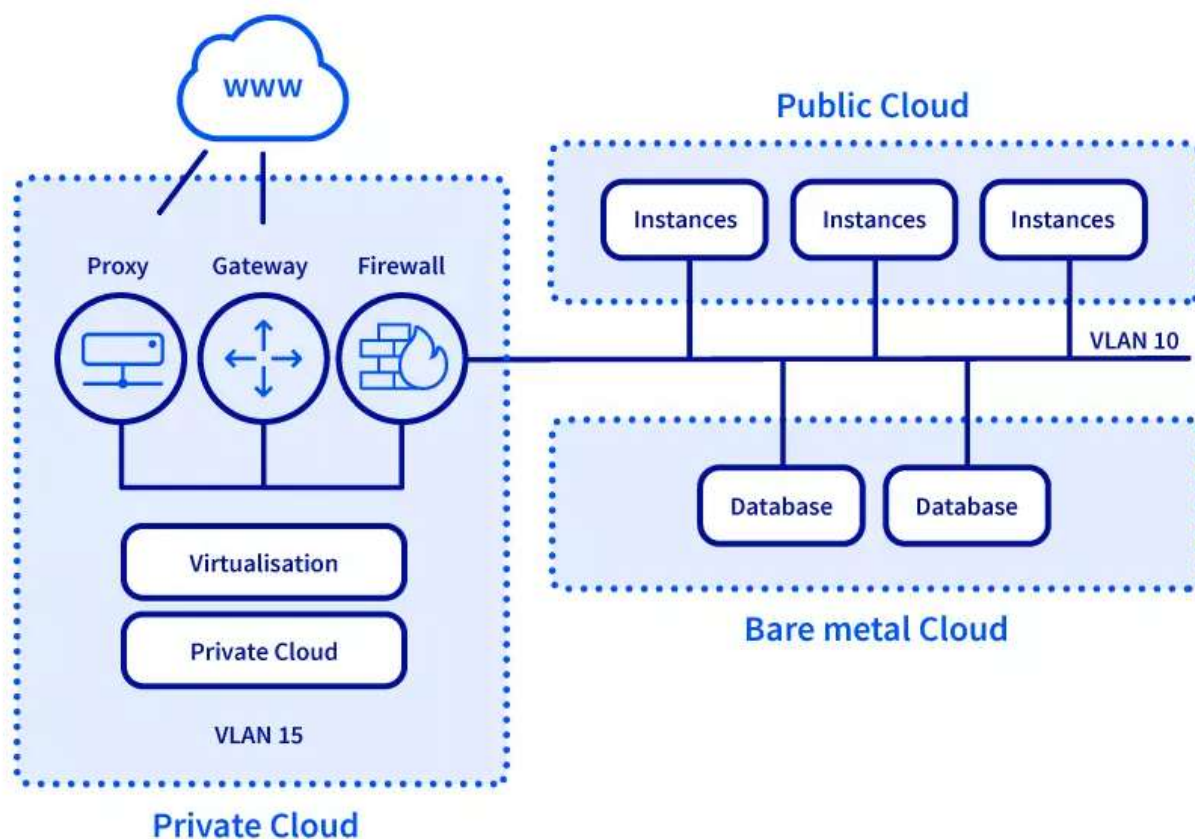
3.2.2. Architecture Applicative

A Compléter avec Infos de la TMA

3.3. vRack

La couche de communication utilisée pour la solution XX'AU repose sur la technologie vRack d'OVHcloud (baie virtuelle). Celle-ci permet de connecter virtuellement plusieurs serveurs entre eux (physiques ou virtuels) en utilisant un réseau privé. Les données inter-serveurs ne transitent pas par le réseau public mais via le réseau OVH physique qui permet ainsi une interconnexion horizontale dédiée et totalement sécurisée. L'ensemble de ce réseau vRACK est sous responsabilité OVH et est complètement géré par celui-ci.

La bande passante théorique du réseau privé vRack est de 10Gbit/s.



Principe de fonctionnement du vRack OVH

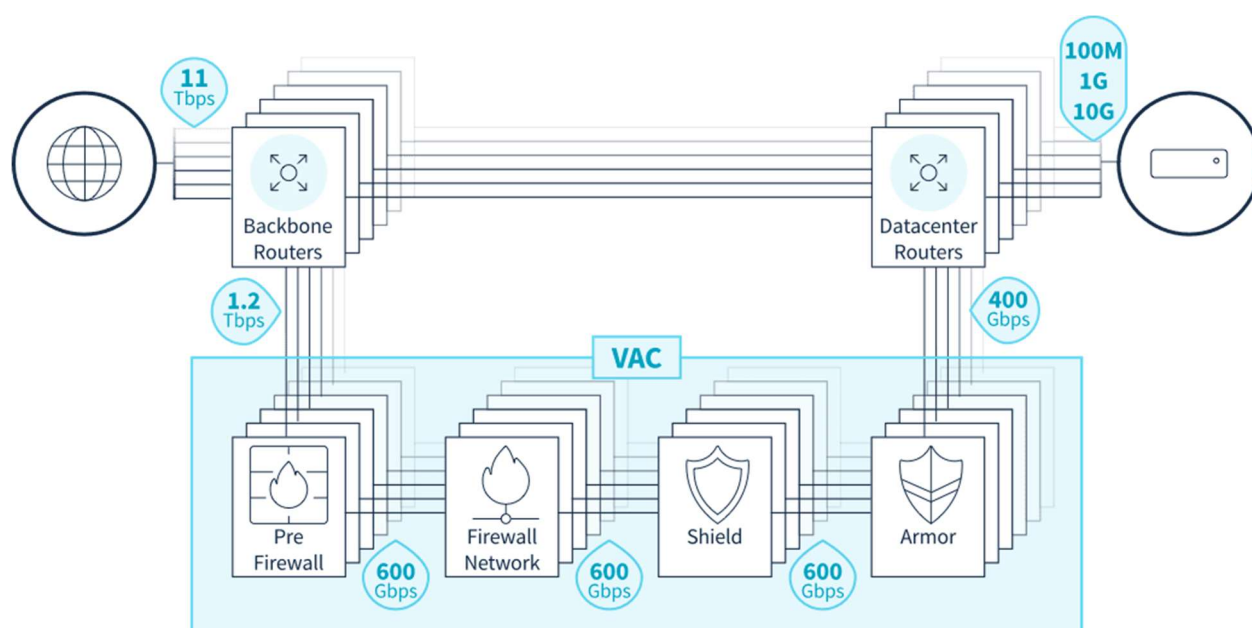
3.4. Interconnexions Réseaux

3.4.1. Protection Anti-DDoS

Afin de garantir une sécurité maximale de l'infrastructure, OVH fournit un système anti-DDoS. Celui-ci s'appuie sur le partenariat technologique avec la société Tileria (force de calcul pour l'intelligence du réseau, l'analyse des flux et le traitement de la sécurité) ainsi que sur la solution « Peakflow » de la société Arbor Networks qui se charge de la détection des attaques et de la mise en place de capacités d'atténuation.

La combinaison de ces deux technologies qui analysent en temps réel et à haute vitesse tous les paquets, absorbent le trafic entrant sur chaque serveur et mitigent en identifiant les paquets IP non légitimes, assure la continuité de service de l'application en cas d'attaque par déni de service.

Cette solution est Native et incluse à tous les services d'OVH qui ont un accès à Internet. Cependant, il n'est pas possible de configurer ce service. OVH reste le propriétaire de cette solution, et est le seul à le manager. Toute l'infrastructure (donc tous les environnements) qui supporte le projet « MTE – XX'AU » est protégée par le service anti-DDoS d'OVH.



Principe de fonctionnement de la protection anti-DDoS d'OVH

3.4.2. Accès Internet

L'accès au réseau Internet est fourni par OVH par l'intermédiaire de la solution HPC SNC sur laquelle le projet « MTE – XX'AU » est hébergé. La bande passante théorique de cet accès Internet est de 10Gbit/s descendant et 1Gbit/s montant. L'accès à Internet ou non des machines depuis le réseau privé se fait via l'implémentation d'une règle de pare feu sur le (ou les) Cluster NSX Edge se situant en tête de la chaîne de connectivité réseau.

3.4.3. Zones Réseaux

Afin de faire communiquer les différents composants du projet « MTE – XX'AU » ainsi que dans le but de les isoler, différents VLAN ont été définis. Des IPs Publiques ont aussi été attribuée au projet sur chaque région OVH.

- Environnement « Production » (SBG) :
 - Réseau Privé : 10.20.13.0/25 - PLT-SBG-PRD-FRONT-OCP
 - IPs Publiques : 178.32.128.144/28
- Environnement « Hors-Production » (RBX) :
 - Réseau Privé : 10.0.0.0/24 - VLAN-OCP-PLT-RBX - VLAN 1000
 - IPs Publiques : 51.77.182.128/28
- Environnement « Reprise » (RBX) : Voir « Hors-Production »

Les régions RBX et SBG sont connectées via un lien « Inter DC » dédié à la solution.

3.4.4. Routage & Filtrage de la Zone Interne

Le routage et la sécurité de la zone interne en amont du Tenant OpenShift est assurée par un étage de Cluster NSX Edge en HA par environnement ayant le rôle de FW d'activé et portant chacun des IPs Publiques et des URLs à usages différents. En cas de défaillance d'un des nœuds du Cluster, le basculement dynamique de l'un vers l'autre est automatique.

L'accès au service depuis l'extérieur s'appuie sur des enregistrements DNS « gouv.fr » qui sont directement gérés par le client MTE. Ces enregistrements pointent eux-mêmes sur des enregistrements DNS de la zone « tdp.ovh » directement gérée par les équipes de l'infogérant.

Ce mécanisme permet de garantir une plus grande flexibilité en cas de modification d'adressage ou en cas de déclenchement du PRA.

Région	IP Publique	URL	Description
RBX	51.77.182.129	*.qualif.platau.cohesion-territoires.gouv.fr	Environnement de Qualification
RBX	51.77.182.129	*.preprod1.cohesion-territoires.gouv.fr	Environnement de PreProd (Bac à Sable)
RBX	51.77.182.129	*.preprod2.cohesion-territoires.gouv.fr	Environnement de PreProd (Performance)
RBX	51.77.182.129	*.demo.platau.cohesion-territoires.gouv.fr	Environnement de Démonstration
RBX	51.77.182.129	*.recette.platau.cohesion-territoires.gouv.fr	Environnement de Recette
RBX	51.77.182.129	*.demo2.platau.cohesion-territoires.gouv.fr	Environnement de Démonstration
RBX	51.77.182.129	*.platau.dev	
RBX	51.77.182.129	dns.apps.acc.platau.tdp.ovh	Enregistrement DNS TDP vers lequel pointe toutes les URLs de Hors-Production
RBX	51.77.182.129	api.plt-rbx.snc.tdp.ovh	Accès à l'API OpenShift Hors-Production
RBX	51.77.182.129	*.apps.plt-rbx.snc.tdp.ovh	
SBG	178.32.128.145	api.platau.cohesion-territoires.gouv.fr	
SBG	178.32.128.145	avisau.cohesion-territoires.gouv.fr	Environnement de Production
SBG	178.32.128.145	rieau.cohesion-territoires.gouv.fr	Environnement de Production
SBG	178.32.128.145	supervision.platau.cohesion-territoires.gouv.fr	
SBG	178.32.128.145	support.platau.cohesion-territoires.gouv.fr	
SBG	178.32.128.145	telechargement.platau.cohesion-territoires.gouv.fr	
SBG	178.32.128.145	dns.apps.prod.platau.tdp.ovh	Enregistrement DNS TDP vers lequel pointe toutes les URLs de Production
SBG	178.32.128.145	api.plt-sbg.snc.tdp.ovh	Accès à l'API OpenShift Production
SBG	178.32.128.145	*.apps.plt-sbg.snc.tdp.ovh	
RBX	51.77.182.129	api.platau.cohesion-territoires.gouv.fr	Uniquement en cas de déclenchement du PRA
RBX	51.77.182.129	avisau.cohesion-territoires.gouv.fr	
RBX	51.77.182.129	rieau.cohesion-territoires.gouv.fr	
RBX	51.77.182.129	supervision.platau.cohesion-territoires.gouv.fr	
RBX	51.77.182.129	support.platau.cohesion-territoires.gouv.fr	
RBX	51.77.182.129	telechargement.platau.cohesion-territoires.gouv.fr	
RBX	51.77.182.129	dns.apps.prod.platau.tdp.ovh	
RBX	51.77.182.129	api.plt-sbg.snc.tdp.ovh	
RBX	51.77.182.129	*.apps.plt-sbg.snc.tdp.ovh	

3.4.5. Load Balancing

Le Load Balancing des URLs en amont du tenant OpenShift est assuré par un étage de deux Cluster NSX Edge en HA par région ayant le rôle de LB d'actif. Les différents serveurs virtuels et pools nécessaires y sont configurés afin d'acheminer le trafic vers le bon équipement en entrée du tenant OpenShift.

ENVIRONNEMENT	EQUIPEMENT	SERVEUR VIRTUEL	POOL	PROTOCOLE	IP SOURCE	PORT SOURCE	ROLE DESTINATION	IP DESTINATION	PORT DESTINATION	COMMENTAIRE
PRODUCTION	NSPLTSBGPRDGTW001	PLT-SBG-PRD-FRONT-OCF-API	PLT-SBG-PRD-FRONT-OCF-API	HTTPS	178.32.128.145	6643	MASTERS	10.20.13.10	6443	
		PLT-SBG-PRD-FRONT-OCF-ONG-HTTPS	PLT-SBG-PRD-FRONT-OCF-APPS	HTTPS	178.32.128.145	443	INFRAS	TOUTES	443	
		PLT-SBG-PRD-FRONT-OCF-ONG-HTTP	PLT-SBG-PRD-FRONT-OCF-APPS	HTTP	178.32.128.145	80	INFRAS	TOUTES	80	
HORS-PRODUCTION	NSPLTRBXACCGTW001	OCF-API	OCF-API-POOL	HTTPS	51.77.182.129	6443	MASTERS	10.0.0.10	6443	
		OCF-APPS	OCF-APPS-POOL	HTTPS	51.77.182.129	443	INFRAS	10.0.0.11	443	
		OCF-APPS-HTTP	OCF-APPS-HTTP	HTTPS	51.77.182.129	80	WORKERS	TOUTES	80	
REPRISE	NSPLTRBXACCGTW001	OCF-API	OCF-API-POOL	HTTPS	178.32.128.145	6643	MASTERS	10.0.0.10	6443	PRA
		OCF-APPS	OCF-APPS-POOL	HTTPS	178.32.128.145	443	INFRAS	10.0.0.11	443	
		OCF-APPS-HTTP	OCF-APPS-HTTP	HTTP	178.32.128.145	80	WORKERS	TOUTES	80	

Le trafic une fois acheminé jusqu'au tenant OpenShift est ensuite traitée par des routeurs Ingress déployés directement dans ce celui-ci.

3.4.6. NAT

Ci-après un récapitulatif des règles de NAT implémentées sur les différents Cluster NSX Edge positionnés en tête de réseau.

ENVIRONNEMENT	EQUIPEMENT	ACTION	ORIGINAL					TRADUIT	
			PROTOCOLE	IP SOURCE	PORT SOURCE	IP DESTINATION	PORT DESTINATION	IP	PLAGE PORT
PRODUCTION	NSPLTSBGPRDGTW001	SNAT	ANY	10.20.13.0/25	ANY	ANY	ANY	178.32.128.145	ANY
HORS-PRODUCTION	NSPLTRBXACCGTW001	SNAT	ANY	10.0.0.0/24	ANY	ANY	ANY	51.77.192.129	ANY
REPRISE	NSPLTRBXACCGTW001	SNAT	ANY	10.0.0.0/24	ANY	ANY	ANY	51.77.192.129	ANY

3.4.7. Pare-Feu

Ci-après un récapitulatif des règles de flux implémentées sur les différents Cluster NSX Edge positionnés en tête de réseau.

ENVIRONNEMENT	EQUIPEMENT	NOM	SOURCE	DESTINATION	SERVICE	PORT SOURCE	ACTION	JOURNAL	COMMENTAIRE
PRODUCTION	NSPLTSBGPRDGTW001	OCP API/APPS	ANY	149.202.13.146	TCP/80 TCP/443 TCP/6443	ANY	ACCEPTER	NON	
		OCP EXT	10.20.13.0/25	ANY	ANY	ANY	ACCEPTER	NON	
		DEFAULT RULE	ANY	ANY	ANY	ANY	REFUSER	NON	
HORS-PRODUCTION	NSPLTRBXACCGTW001	OCP API/APPS	ANY	51.77.182.129	TCP/80 TCP/443 TCP/6443	ANY	ACCEPTER	NON	
		OCP EXT	10.0.0.0/24	ANY	ANY	ANY	ACCEPTER	NON	
		DEFAULT RULE	ANY	ANY	ANY	ANY	REFUSER	NON	
REPRISE	NSPLTRBXACCGTW001	OCP API/APPS	ANY	51.77.182.129	TCP/80 TCP/443 TCP/6443	ANY	ACCEPTER	NON	PRA
		OCP EXT	10.0.0.0/24	ANY	ANY	ANY	ACCEPTER	NON	
		DEFAULT RULE	ANY	ANY	ANY	ANY	REFUSER	NON	

3.5. Stockage Objet

L'offre de stockage objet « Public Cloud » d'OVH est utilisée afin de stocker les milliers de fichiers de données liées au projet « MTE – XX'AU ». Ces Buckets de stockage sont tous déployés avec le type « Privé » et leur accès n'est possible que via l'utilisation d'une clé d'accès, et d'une compte utilisateur.

Elles sont hébergées par défaut sur la région de Gravelines, et répliquées en permanence sur la région de Strasbourg dans le cadre du PRA. En cas de déclenchement du PRA le sens de synchronisation est temporairement inversé jusqu'à revenir en fonctionnement nominal soit une Production sur Strasbourg.

Un projet y est dédié à la Production (PRD-MTES-PLATEAU) et un second est dédié à tous les autres environnements (ACC-MTES-PLATEAU).

3.6. Sauvegarde

3.6.1. Kasten

Les Workloads Applicatifs (Applications) et leurs données persistantes (PVC) de l'infrastructure OpenShift sont sauvegardées via la solution Kasten K10 (Veeam). Une consistance des données applicatives sauvegardées est prise en compte grâce à des Blueprints Kanister. Ceux-ci permettent de pouvoir effectuer des actions avant, après et/ou pendant la sauvegarde d'une application.

« Kasten » est installé sur les différents Cluster Kubernetes afin d'en sauvegarder les containers. Kasten s'occupe d'effectuer la sauvegarde des applications situées sur les différents Workers déployés pour le projet.

Les données ainsi sauvegardées sont envoyées dans une zone de stockage de type Object Storage High Performance - S3 dédiée au projet « MTE – XX'AU ». Ces sauvegardes sont chiffrées et stockées sur la région OVH de Gravelines (GRA).

Le plan de sauvegardes Kasten pour le projet est le suivant :

Environnement	Périodicité	Rétention
PRODUCTION	Daily Exported Snapshots	7
	Weekly Exported Snapshots	4
	Monthly Exported Snapshots	12
	Yearly Exported Snapshots	5
PRE-PRODUCTION	Daily Exported Snapshots	7
	Weekly Exported Snapshots	4
	Monthly Exported Snapshots	12
DEMONSTRATEURS	Daily Exported Snapshots	3
	Weekly Exported Snapshots	4
QUALIFICATIONS	Daily Exported Snapshots	3
RECETTE	Daily Exported Snapshots	3

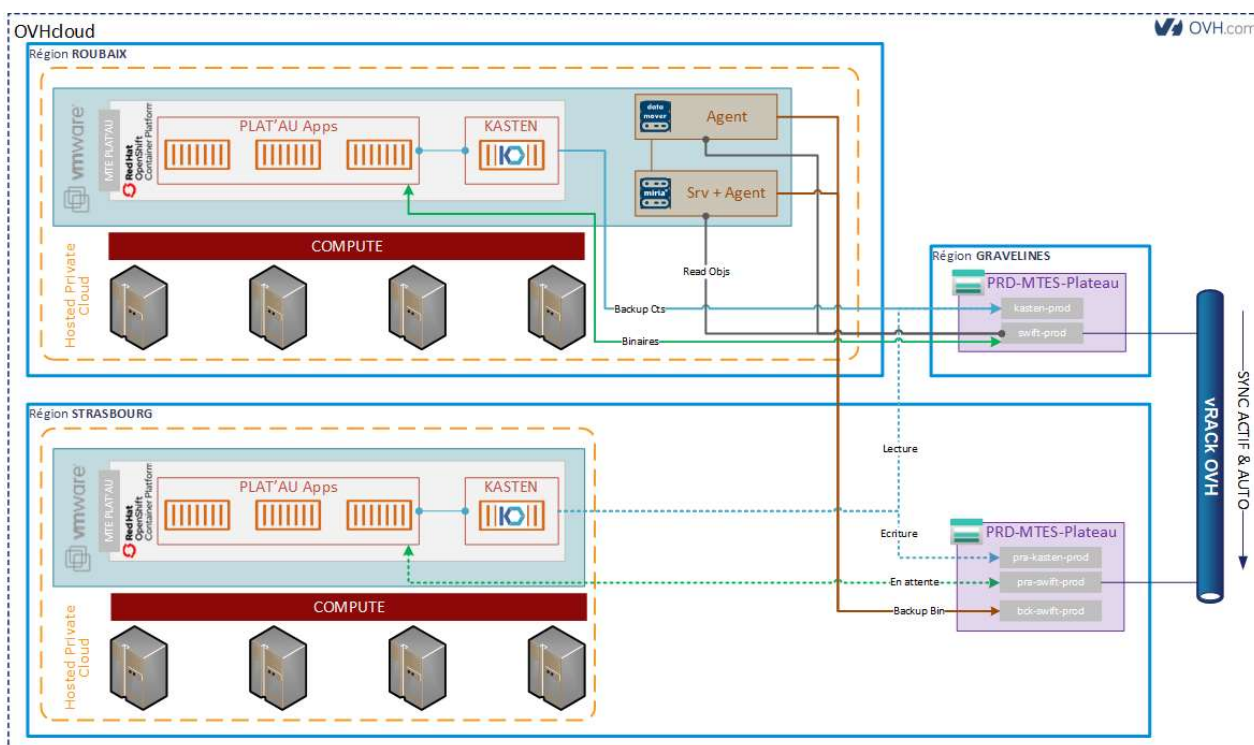
Par ailleurs l'environnement de **Production est sauvegardé toutes les 4h.**

3.6.2. MIRIA

MIRIA de l'éditeur ATEMPO est une solution pour la sauvegarde, la migration, la synchronisation et le déplacement de fichiers non structurés entre stockages hétérogènes de grandes volumétries.

Elle permet la sauvegarde des différents environnements SWIFT attachés au projet « MTE – XX'AU » qui de par leur volumétrie importante et leur très grand nombre de fichiers ne pouvaient faire l'objet d'une sauvegarde plus classique.

La solution est installée sur des VMs de type IaaS déployées sur la région de Strasbourg. En cas d'activation du PRA, le service est basculé sur des VMs hébergées sur la région de Roubaix.



Principe de fonctionnement de la solution de sauvegarde MIRIA

Les données sauvegardées sont envoyées dans une zone de stockage de type Object Storage High Performance - S3 dédiée au projet « MTE – XX'AU » sur la région de Strasbourg.

Les sauvegardes MIRIA ne sont réalisées que sur l'environnement de Production et interviennent toutes les 4h à partir de minuit chaque jour.

3.7. DNS

Dans le cadre du projet « MTE – XX'AU », certaines URLs sont portées par la zone DNS « tdp.ovh » propre à l'infogérant. Ci-après un rappel des entrées concernées :

DOMAINE	TTL	TYPE	CIBLE
dns.apps.acc.platau.tdp.ovh.	0	A	51.77.182.129
dns.apps.prod.platau.tdp.ovh.	0	A	178.32.128.145
kibana.apps.prod.platau.tdp.ovh.	0	A	178.32.128.145
api.plt-rbx.snc.tdp.ovh.	300	A	51.77.182.129
api.plt-sbg.snc.tdp.ovh.	300	A	178.32.128.145
*.apps.plt-rbx.snc.tdp.ovh.	300	A	51.77.182.129
*.apps.plt-sbg.snc.tdp.ovh.	300	A	178.32.128.145

3.8. Service SMTP

Le service Private Exchange 2019 d'OVH a été choisi pour répondre au besoin de relais mail dans un premier temps (envois de notifications) et à des fins de réceptions de messages ultérieurement.

La zone DNS suivante « platau.cohesion-territoires.gouv.fr » a été créée par MTES et associée au service de messagerie. La boîte aux lettres admin@platau.cohesion-territoires.gouv.fr permet de gérer le système qui est accessible via Webmail :

<https://mail.platau.cohesion-territoires.gouv.fr>

Note : la configuration se fait depuis la console OVH (Section « Web »). Les identifiants étant stockés par l'infogérant.

4. OpenShift

4.1.1. Présentation

La couche de conteneurisation est amenée par la plateforme OpenShift en tant que service (PaaS) de Red Hat qui est basée sur des conteneurs Linux au format Docker, l'orchestration Kubernetes (K8S) et Red Hat Enterprise Linux (RHEL)

Kubernetes apporte les fonctions clés de l'orchestration de Conteneurs et une grande robustesse :

- Automatisation du déploiement : Kubernetes permet de modéliser une application sous conteneurs et d'en automatiser le déploiement en « Rolling update » sans coupure d'accès à cette application. On peut aussi redéploier en quelques clics un environnement ou une nouvelle application en s'appuyant sur les modèles déjà créés.
- Gestion de Clusters : Kubernetes apporte des outils puissants et robustes pour gérer le Cluster, comme l'Ingres qui joue un rôle de LoadBalancer Nginx très fiable, ou la gestion intelligente des flux.
- Haute disponibilité : En mode multi-managers et multi-nœuds (les « Workers »), Kubernetes est un cluster en haute disponibilité qui sait monitorer l'état des conteneurs et relancer automatiquement les services applicatifs en redéploiant les conteneurs.
- Scalabilité : Kubernetes apporte une double scalabilité en permettant l'ajout de conteneurs tout comme de nœuds « Workers » dans le Cluster ; il sait gérer un volume très élevé de conteneurs et apportera une grande scalabilité aux applications.

La couche OpenShift vient compléter Kubernetes avec une interface qui permet d'en exploiter au mieux les fonctionnalités :

- Interface graphique : OpenShift apporte une vue graphique aux conteneurs et simplifie l'usage de certaines fonctionnalités de Kubernetes.
- Gestion de « projets » et gestion de droits : OpenShift apporte un modèle d'organisation plus complet, il est possible de créer des « projets » et gérer des utilisateurs, avec une gestion différenciée des droits par projet et par utilisateur.
- Simplifie et automatise les versions d'applications et de conteneurs, les déploiements, la mise à l'échelle, la gestion de l'intégrité.

Le déploiement automatisé d'OpenShift est effectué via un Runner GitLab hébergé chaque région OVH.

4.1.2. Certificats

OpenShift utilise des certificats signés afin d'authentifier les différentes URLs. Les certificats étant sous responsabilité MTE, l'infogérant intervient dans la génération du CSR et la fourniture à MTE ainsi que dans l'implémentation des certificats sur les environnements.

L'expiration des certificats est supervisée avec génération automatique d'un ticket un mois avant la date d'expiration effective.

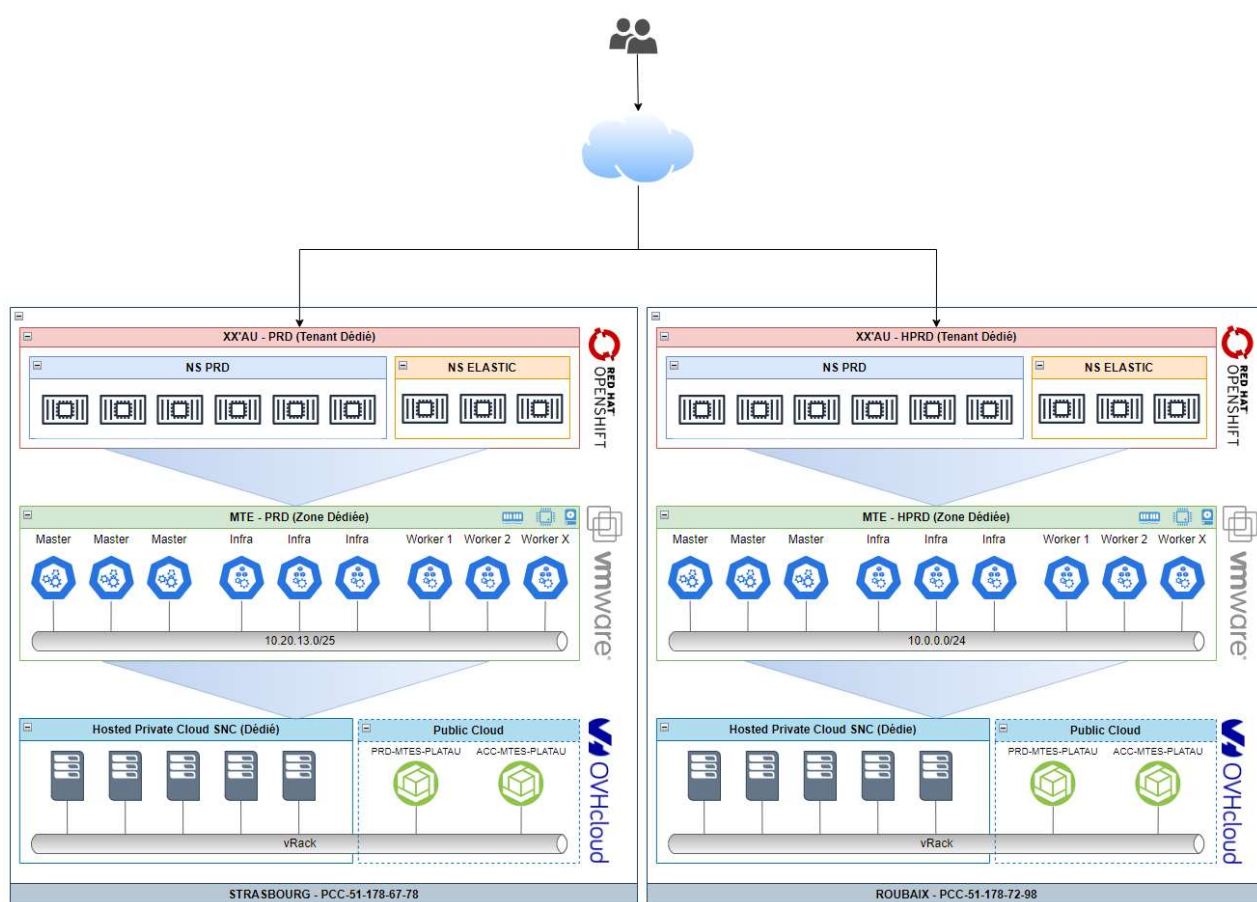
4.1.3. Environnements

Deux environnements de Production et Hors-Production hébergeant les différents NameSpace applicatifs sont déployés sur deux infrastructures dédiés sur deux régions différentes. Ils contiennent eux-mêmes les différents environnements Applicatifs liés au projet « MTE – XX'AU ». En cas de déclenchement du PRA, l'environnement de Production sera temporairement hébergé sur l'infrastructure de Hors-Production. Les différents environnements de Hors-Production seront éteints pendant toute la durée d'activation du PRA.

Chaque environnement HPC SNC dispose de ses propres VMs et de ses propres ressources en CPU, RAM et stockage. Les différents NameSpaces présents dans chaque environnement utilisent de manière mutualisée les ressources qui leurs sont mises à disposition.

ENVIRONNEMENT	ROLE	NOMBRE	GABARIT	CPU	RAM	TOTAL CPU	TOTAL RAM
PRODUCTION	MASTER	3	X1.XLARGE	4	16	12	48
	INFRA	3	X1.XLARGE	4	16	12	48
	WORKER	9	N/A	24	64	216	576
TOTAL						240	672
HORS-PRODUCTION	MASTER	3	X1.XLARGE	4	16	12	48
	INFRA	3	X1.XLARGE	4	16	12	48
	WORKER (PPRD)	6	N/A	24	64	144	384
	WORKER (ACC)	4	C1.2XLARGE	12	32	48	128
TOTAL						216	608

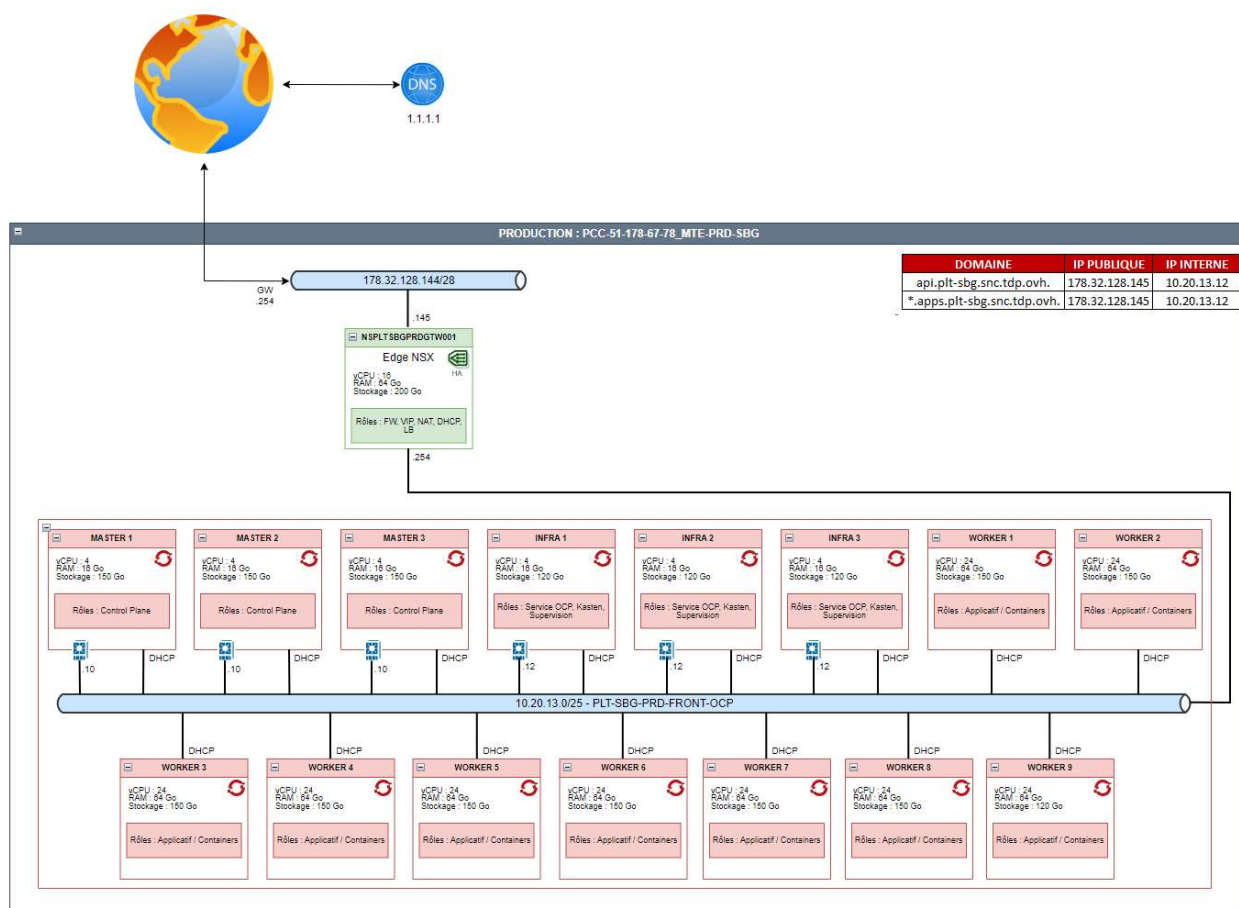
a. Architecture MTE XX'AU (HLD du Socle Technique)



Design « High Level » de principe de la solution XX'AU

b. Environnement de Production

L'environnement OpenShift de Production est hébergé sur la région OVH de SBG, il n'héberge qu'un environnement Applicatif.



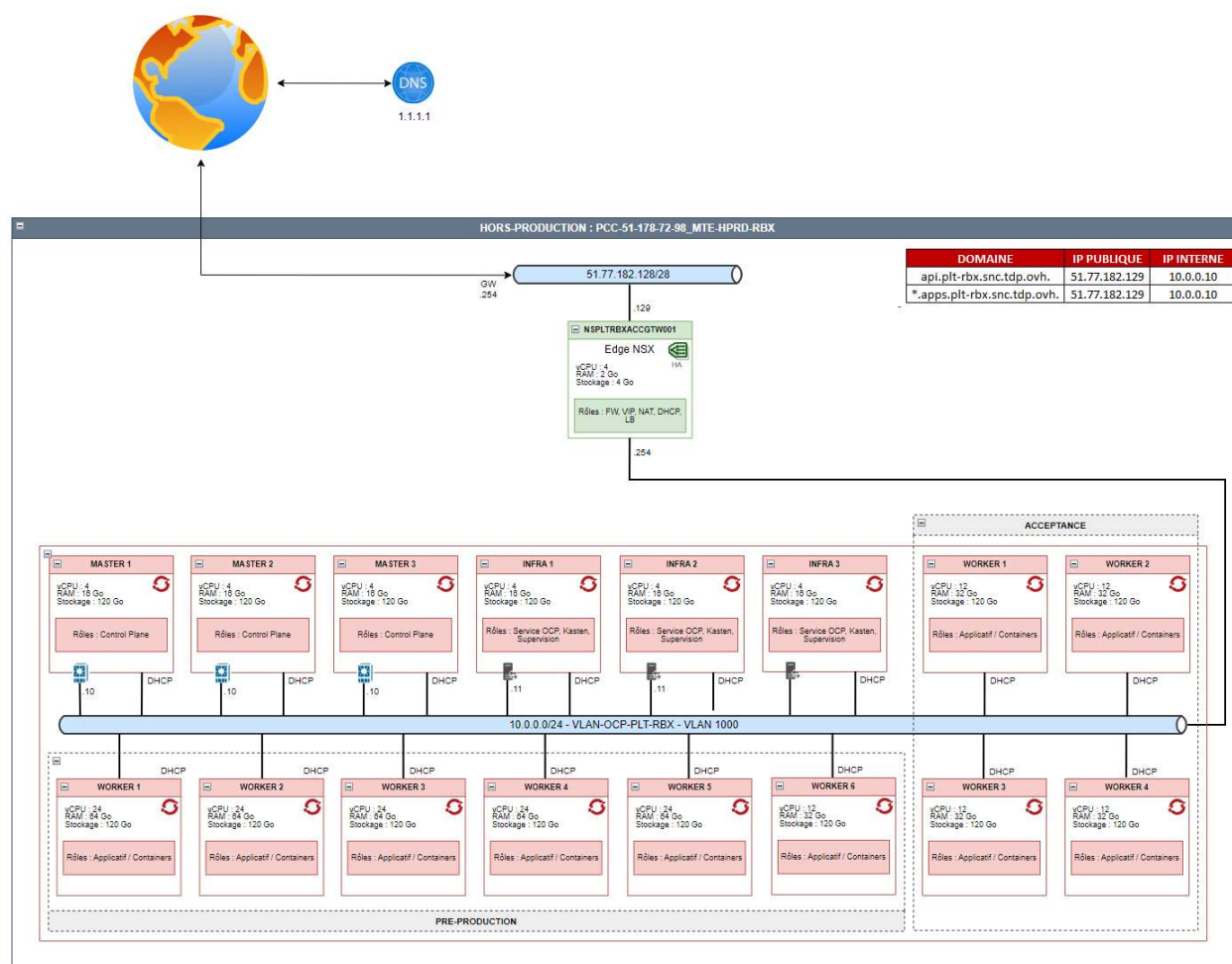
Focus de l'environnement de Production

c. Environnement de Hors-Production

L'environnement OpenShift de Hors-Production est hébergé sur la région OVH de RBX, il héberge plusieurs environnement Applicatifs :

- PreProd
- Recette
- Démonstrateur

Des Workers sont dédiés à la PreProd, d'autres Workers (dits d'Acceptance) sont dédiés aux autres environnements de Hors-Production.



Focus de l'environnement de Hors-Production

d. Environnement de Reprise

L'environnement OpenShift de Reprise (PRA) est hébergé sur la région OVH de RBX, il n'héberge qu'un environnement Applicatif correspondant à l'environnement de Production. En cas de déclenchement du PRA les environnements de Hors-Production sont arrêtés et l'environnement de Production est restauré sur le Cluster OpenShift de Hors-Production. Une documentation dédiée de type RFC existe spécifiquement pour traiter de cette bascule.

5. Plan Reprise Activité

Un plan de reprise d'activité (PRA) est implémenté pour l'environnement de Production du projet « MTE – XX'AU » notamment par utilisation de la solution MIRIA pour la réplication continue des milliers de fichiers de données liés au projet. La restauration des données applicatives étant par ailleurs traitée via les sauvegardes Kasten.

Les actions inhérentes à la bascule en cas de déclenchement du PRA sont décrites dans le document « PRA - TDP MTE PLAT'AU.docx ».

6. Supervision

6.1. Prometheus / Grafana

Prometheus est une solution open-source de « monitoring » et « d'alerting ». L'objectif de Prometheus est de fournir, ici, un service de supervision permettant de surveiller dans son intégralité la couche de l'infrastructure OpenShift ; c'est-à-dire les informations relatives aux machines hôtes de l'OpenShift ainsi que des instances virtuelles portées par ce dernier (RAM, CPU, Network...) et de leurs principaux composants clefs.

L'adjonction de la solution « Grafana » permet de fournir un affichage graphique de ces différentes informations.

Des processus de collecte de métriques (Exporter) sont implémentés sur l'ensemble des machines « XX'AU ». Ces Exporters sont interrogés via API par Prometheus qui communique à Grafana ces informations pour mise en forme graphique.

Un système d'alerte permet d'informer l'exploitant en cas de détection d'incidents.



Exemple de Dashboard Grafana

6.2. AlertManager

Le service RedHat AlertManager, apporte un service d'envoi de message d'alerte aux utilisateurs sur ordre du serveur de supervision Prometheus. Il embarque des fonctionnalités d'agrégations, de duplication de messages pour limiter le spam et de routage d'alertes en fonction du type d'alerte ainsi que du destinataire concerné.

6.3. Dynatrace

La supervision de la solution est réalisée à l'aide de la solution Dynatrace qui se présente sous la forme d'une plateforme, unique, dédiée à l'observabilité et à l'analyse de l'ensemble de la chaîne applicative. Ce monitoring intégral dit « Full Stack », vise à donner une vision exhaustive de l'état des performances des applications mais aussi des infrastructures y compris des infrastructures conteneurisées.

6.3.1. Modes de surveillance

Deux modes de surveillance sont possibles : « Infrastructure Monitoring » et « Application Performance Monitoring » :

- Monitoring de l'infrastructure uniquement « **Infra Only** » : ce mode simplifié et automatisé donne de la visibilité en continu sur les hôtes, les machines virtuelles, les conteneurs, la partie réseau ainsi que tous ce qui est relatif aux événements et journaux.
- Surveillance des performances des applications « **APM** » : une visibilité automatisée au niveau du code et des transactions est fournie pour l'ensemble de la chaîne applicative. Ici, Dynatrace détecte les interdépendances, les anomalies et indique les causes profondes d'un problème. Les micro-services sont également détectés et monitorés automatiquement.

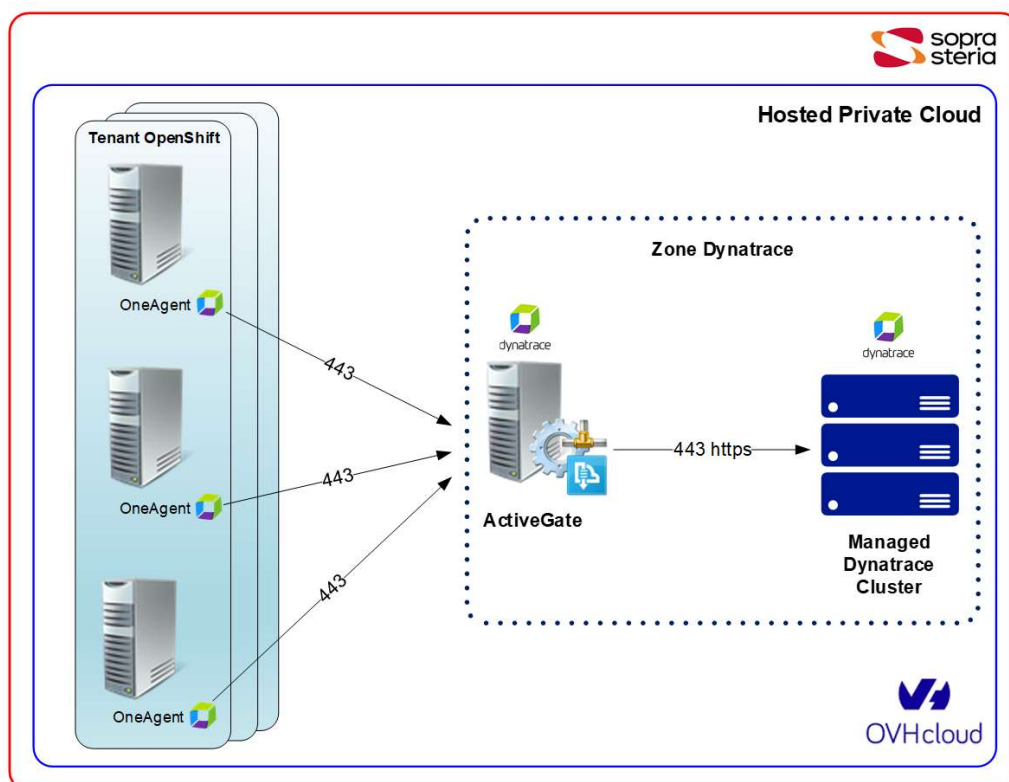
Sur le projet « MTE – XX'AU » les deux modes de surveillance sont implémentés.

6.3.2. ActiveGate

Dynatrace ActiveGate est un proxy dédié qui se place entre les agents Dynatrace OneAgent et le serveur Dynatrace Managed (Cluster Central). Dans le cadre d'un grand nombre de connexions sortantes et d'une quantité de trafic importante, celui-ci a pour rôle de canaliser et d'acheminer le trafic d'information remonté par les agents « OneAgent » vers le serveur Dynatrace Managed central.

6.3.3. Principe de fonctionnement de la supervision Dynatrace

A chaque minute, Dynatrace interroge les plugins ActiveGate. Chaque plugin interroge à son tour l'application qu'il doit surveiller (via https/443), en détermine l'état et renvoie l'information vers le serveur Dynatrace Managed Server (via https/443) sous la forme de métriques.



Principe de fonctionnement de la solution Dynatrace.

7. Opsgenie : Gestion des Alertes et Incidents

La gestion des alertes et des incidents est confiée à la plate-forme de « résolution continue » Opsgenie de l'éditeur Atlassian. La force de cette solution est de permettre aux équipes DevOps et IT Ops de rester connectées et de collaborer en temps réel tout en conservant une maîtrise des incidents durant tout leur cycle de vie.

Opsgenie prend en charge les intégrations natives pour plus de 200 applications, dont Datadog, Jira, Gitlab, Jenkins, Prometheus, Dynatrace, etc... L'intégration d'autres applications sont rendues possibles en utilisant les API REST ou API de messagerie d'OpsGenie. OpsGenie s'intègre aux outils et services de surveillance et garantit que les personnes clés sont informées (via e-mail, SMS, appels téléphoniques...).

Les outils de gestion des alertes et des appels ont été personnalisés par l'infogérant afin de faciliter le contrôle granulaire des préférences de notification pour les services, les équipes et les utilisateurs individuels (via la configuration d'un mécanisme d'auto-escalade, de définition de plages horaires d'appels, de règles de routage et d'escalades des équipes de support).

The screenshot displays the Opsgenie incident management interface. At the top, the incident is identified by ID #41 and a status of 'P1' (Priority 1). The title is '[Dynatrace] 31 "HTTP monitor global outage" occurred with impact level'. The incident is marked as 'Closed'. The header also shows the impact duration and elapsed time, both at 0H 21M 9S.

The main content area is divided into several sections:

- Postmortem:** A section for the post-mortem report, with a link to 'Assign a due date' and a 'PUBLISHED' status.
- Description:** A section for the incident description, with a link to 'Give more details about this incident'.
- Impacted services:** A section showing the impacted services, with a link to 'Add impacted service'.
- Potential causes:** A section for investigating potential causes, with a link to 'Investigate'.
- Extra properties:** A section for additional properties, with a link to 'Add extra property'.

The timeline on the right side of the interface shows a sequence of events:

- 15:31 - Responder alert acknowledged - API
- 15:31 - Incident resolved - [Redacted]
- 15:31 - Stakeholders updated - System
- 15:31 - The problem was solved [Nginx service was down and it was started successfully. Our team will work on an automated process to start the service when it will be down again in order to avoid this incident]
- 15:39 - Responder alert acknowledged - [Redacted]
- 15:39 - [Redacted]
- 15:39 - Checked on the VM, nginx is down
- 15:38 - Stakeholders updated - System
- 15:38 - We will keep you posted as soon as the incident is solved |
- 15:38 - Stakeholders updated - System
- 15:38 - We are working on the incident global outage |
- 15:35 - Command center session started - System
- Renault incident room

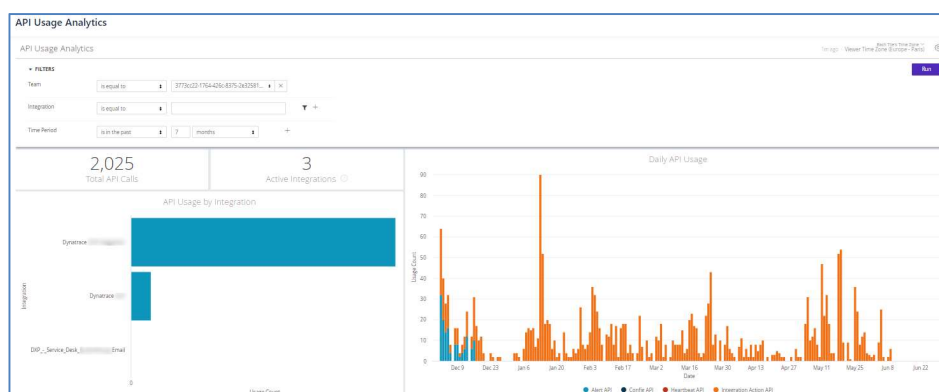
Exemple d'incident avec détails



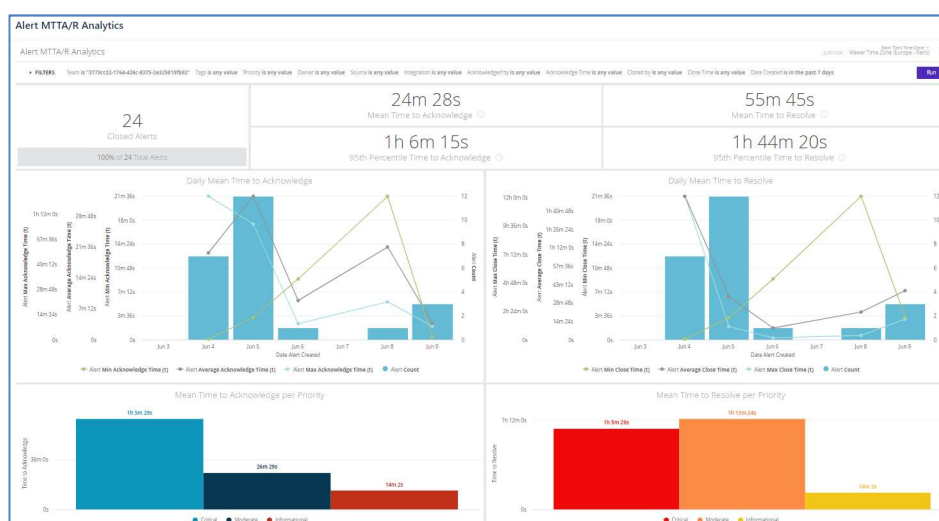
Temps moyen de résolution des alertes

Grâce à sa partie d'analyse, il offre plusieurs « Dashboards » permettant de suivre les données des alertes.

L'analyse sur l'usage des intégration et API (ex : La source d'où proviennent le plus d'alertes (Top API call) :



Le temps moyen de résolution des alertes et incidents par priorité et par jour de la semaine :



L'analyse sur les alertes (ex : nombre des alertes fermées par rapport à celle ouvertes sur une période :



Tous ces rapports peuvent être téléchargés ou envoyés par e-mail avec différents formats CSV, PDF... aux responsables des projets. De plus l'envoi de ces rapports peut être planifié quotidiennement, hebdomadairement.

8. Fin de document
