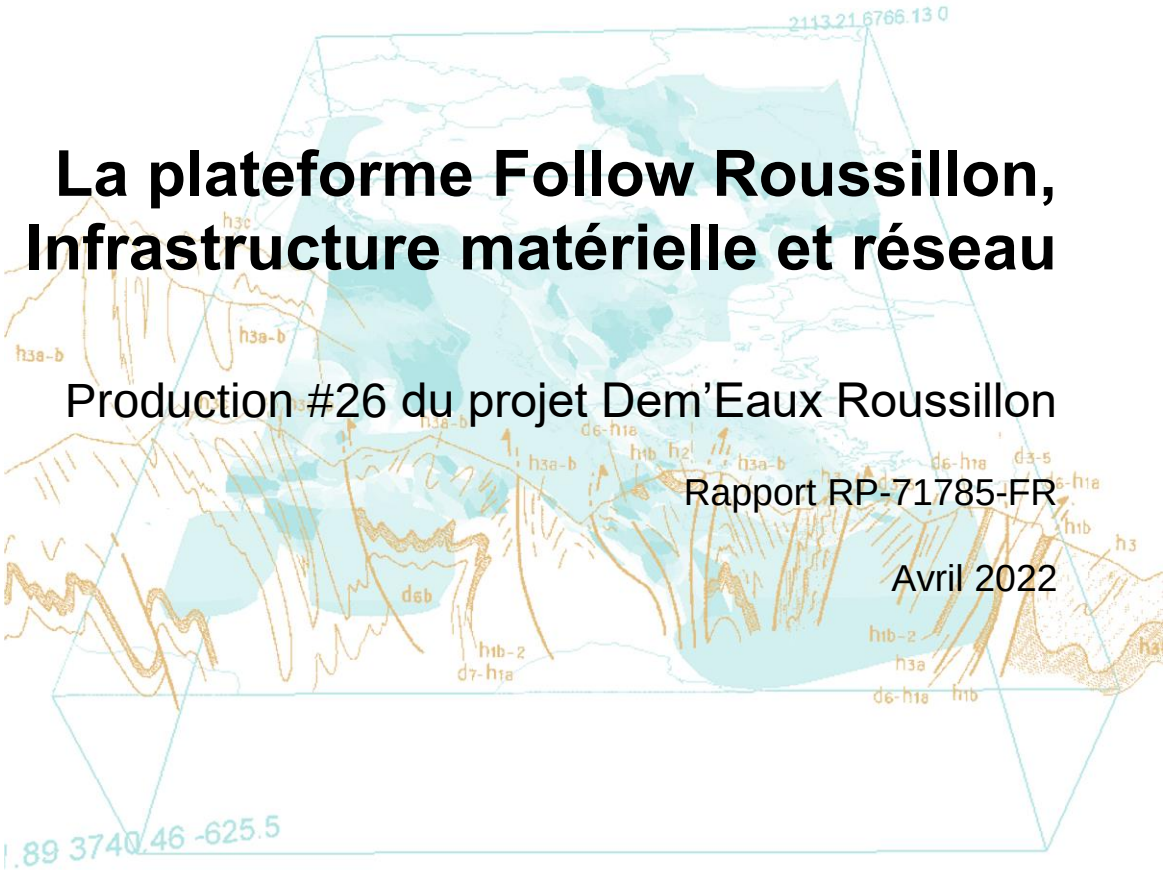
Document public

# La plateforme Follow Roussillon, Infrastructure matérielle et réseau



## Production #26 du projet Dem'Eaux Roussillon

Rapport RP-71785-FR

Avril 2022

Projet cofinancé par le Fonds Européen de Développement Régional

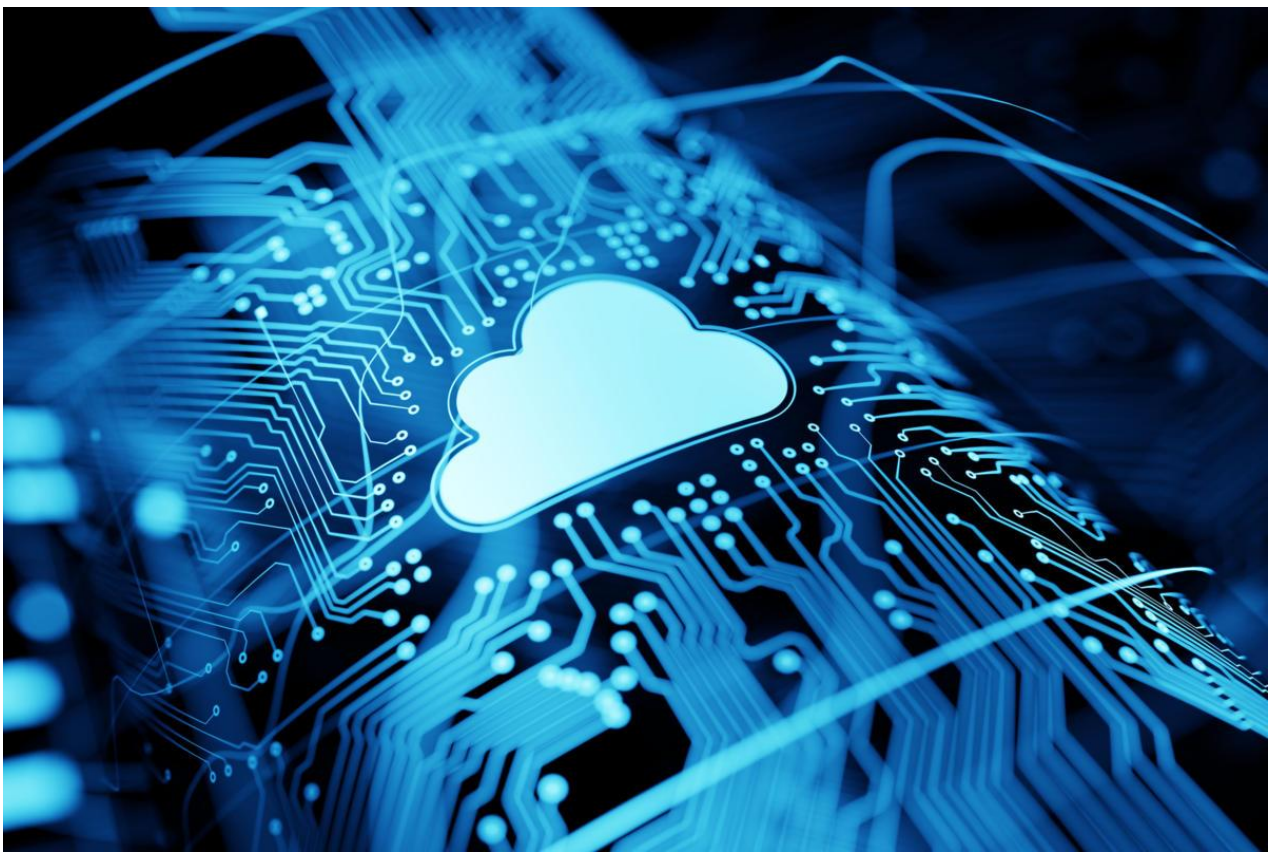
**Mots-clés**: Site internet; Plateforme de valorisation et de gestion, Pyrénées Orientales; Plaine du Roussillon.

En bibliographie, ce rapport sera cité de la façon suivante : Raynaud, J.-B. (2022), La plateforme Follow Roussillon, Infrastructure matérielle et réseau. Production #26 du projet Dem'Eaux Roussillon. Rapport RP-71785-FR, 31p.

synapse

# PROJET DEM'EAUX ROUSSILLON

## *La plateforme Follow Roussillon - Infrastructure matérielle et réseau*



VERSION 1.0

La plateforme Follow Roussillon, Infrastructure matérielle et réseau - Production #26 du projet Dem'Eaux Roussillon

SOMMAIRE



## Révisions

| Version | Description | |
|---|---|---|
| 1.0 | Janvier '22 | - JBR |

## Glossaire

| Terme | Description |
|---|---|
| **1GbE, 10GbE** | Abréviations des débits ethernet de 1 giga bit par seconde et respectivement de 10 gigabits par seconde. |
| **Transit IP, TIP (Cicuit de)** | La connexion réseau d'une installation physique de plateforme dans une baie / rack, avec internet. Elle se matérialise par une arrivée fibre dédiée à la plateforme. |
| **Failover** | Basculement. Terme utilisé avec des services de cluster (à basculement), permettant de basculer les services soutenus par un ou plusieurs composants vers d'autres composants, automatiquement. |
| **NLB** | **Network Load Balancing.** Technologie de répartition de charge au niveau du réseau, supporté nativement par Windows Server. |
| **NAT** | **abr. Network Address Translation,** ou translation des adresses lors du routage des adresses publiques vers des adresses de services internes à la plateforme |
| **hyperconvergence, hyperconvergé** | **Hyperconverged**. Patron d'architecture de clustering permettant la création du niveau de stockage d'un cluster par l'agrégation et la réplication du stockage de chaque nœud du cluster. |

| | |
|---|---|
| **hyper-V** | Le système de virtualisation proposé par Windows Server.<br>Pour NextNet, la couche de virtualisation est assurée par Windows Server 2019 (édition datacenter, édition standard) |
| **endpoint** | Point de terminaison, http, ftp, tcp, ….. En IP un point de terminaison indique une adresse et un port. Un endpoint http est désigné par une URL. |
| **backend** | La partie backoffice d'un système d'information. Les services de collecte ou échange Follow, ou la passerelle (gateway) Iridium sont des modules backoffice |
| **gateway** | Ou passerelle, de backend. La passerelle expose des points de terminaison et permet de router des flux et des messages vers le points de traitements applicatifs . |
| **trunking** | Agrégation des commutateurs réseau pour former un seule commutateur à partir de deux commutateurs distincts. |
| **OS10 (Enterprise Edition)** | Le système d'exploitation de la couche commutation , équipant les commutateurs NextNet |
| **VLT** | abréviation Virtual Link Trunking. Technologie de trunking proposé par Dell dans l'OS 10 équipant les commutateurs S4128F-On utilisés sur NextNet |
| **VRRP** | abrév. Virtual Router Redundancy Protocol. Protocole destiné à augmenter la disponibilité de la passerelle par défaut pour es hôtes d'un même sous-réseau. Supporté sur OS10, équipant les commutateurs NextNet. |
| **VRF** | abrév. Virtual Routing and Forwarding. Technologie de partitionnement d'un routeur physique en plusieurs routeurs virtuels, isolés. Supporté sur OS10, équipant les commutateurs NextNet. |
| **IIS** | **Internet Information Services**, la pile serveur web de Microsoft |
| **ARR (IIS)** | Application Request Routing - Technologie de routage applicatif au sein d'une ferme de serveurs Windows Server, basée sur le routage des demandes externes (urls publiques) vers des points de service applicatifs. |
| **hyperconvergence, hyperconvergé, HCI** | **eng. hyperconverged**. Patron d'architecture proposant l'implémentation du stockage d'un cluster par l'agrégation et la réplication du stockage local de chaque nœud du cluster. Pour ce faire, les nœuds doivent être capable d'échanger / répliquer des données à haut débit réseau, comme 10Gb ou plus. |
| **HCI** | **abrév. Hyperconverged Infrastructure** |
| **S2D** | **Storage Spaces Direct**. La technologie de stockage distribué pour les architectures hyperconvergées, proposée par Microsoft. Elle est combinée en général avec les clustering de basculement (failover) et la virtualisation des nœuds de calcul fonctionnant sur le cluster. |
| **on premise, on prems** | Terme désignant les déploiements "**sur le site" et l'infrastructure du** gestionnaire; par rapport aux déploiement des applications sur des infrastructures cloud opérées par des tiers. |
| **SaaS** | **Software as a Service.** Déploiement externalisé des applications, depuis une perspective utilisateur. Les services de concentration Follow supporte le déploiement SaaS. |
| **TTL** | **Time To Live**. Paramètre de configuration DNS spécifiant la durée de vie d'une association nom-adresse ip. Les valeurs longues permettent à un client DNS d'utiliser son propre cache DNS avant de solliciter à nouveau le serveur DNS, à l'expiration du TTL, les valeurs réduites provoquent des nouvelles requêtes DNS. Dans notre contexte, un TTL réduit comme 5 minutes ou 1 minute permet la mise en place des mécanismes de failover DNS |

| | |
|---|---|
| **DFS/R , DFSN** | **Distributed File Service, Distributed File Service Replication, Distributed File Service Namespaces**. La technologie de distribution, réplication et adressage des espaces de stockage fichier proposé dans Windows Server. DFSN permet d'adresser des espaces de stockage à travers un système de nom. DFS utilise les services Active Directory |
| **Hyper-V** | Le système de virtualisation proposé par Windows Server.<br>Pour NextNet, les hôtes de la virtualisation sont Windows Server 2019, les éditions datacenter et standard. |
| **endpoint** | Point de terminaison serveur pour un protocole comme **http, ftp, tcp, udp, etc)**. Un point de terminaison tcp ou udp indique une adresse ip et un port. Un endpoint http est désigné par une URL. |
| **backend** | Ou backoffice d'un système d'information, constitué de services / modules non exposés publiquement à une utilisation interactive, réalisant les fonctions du système d'information. Dans le domaine de la concentration les services de collecte ou d'échange Follow, sont des services de type backend |
| **gateway** | Ou **passerelle**, de backend. La passerelle expose des points de terminaison et permet de router des flux et des messages utilisateur vers les points de terminaison et de traitement prévus par les applications. |
| **IIS** | **Internet Information Services**, la pile serveur web de Microsoft |
| **ARR (IIS)** | Application Request Routing - Technologie de routage applicatif au sein d'une ferme de serveurs Windows Server, basée sur le routage des demandes externes (urls publiques) vers des points de service applicatifs. |

# 1 OBJET DU DOCUMENT

Ce document a pour objet de présenter l'architecture technique et applicative de la plateforme numérique hébergeant les services de concentration et de valorisation des données des projets de R&D Dem'Eaux Roussillon et Dem'Eaux Thau. Il est présenté dans un premier temps l'infrastructure réseau de la plateforme et dans un second temps l'architecture applicative.

# 2 ARCHITECTURE RESEAU

## 2.1 Introduction

### 2.1.1 NextNet

NextNet est la plateforme applicative de production pour les services de concentration et de valorisation des données SaaS des projets de R&D Dem'Eaux Thau et Dem'Eaux Roussillon.

Ce document présente l'infrastructure réseau de cette plateforme en terme de:

- équipements utilisés
- la configuration réseau
- les mécanismes et configurations adoptés pour assurer la disponibilité de la plateforme
- les outils d'administration (plans, outillage) utiles pour la gestion et l'administration opérationnelle de la plateforme

### 2.1.2 Références documentaires

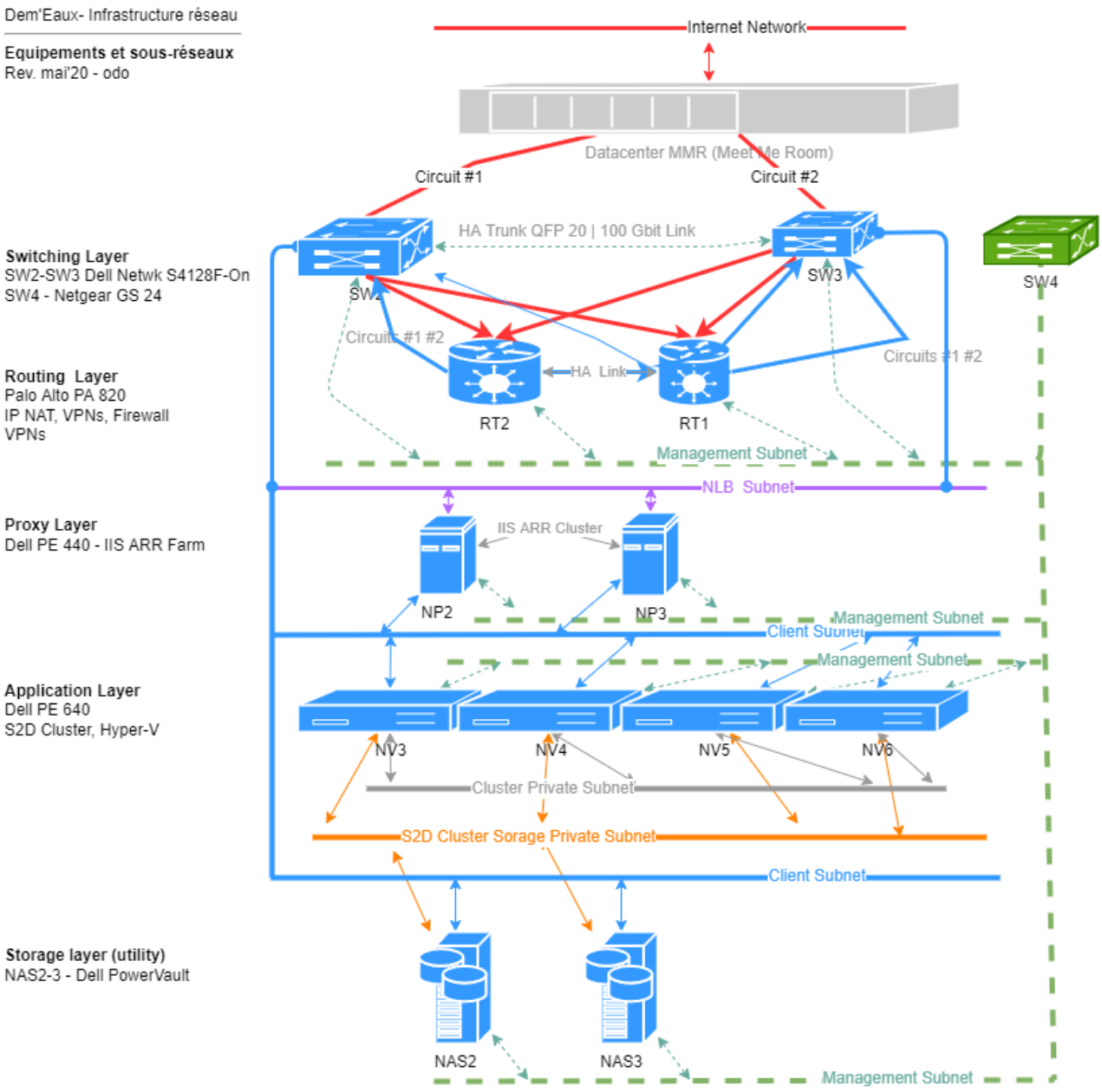
Le tableau suivant liste des références documentaires utiles pour maintenir le document et disposer de précisions supplémentaires.

| Abréviation | Description | Commentaire |
|---|---|---|
| | ***Références fournisseurs / constructeurs*** | |
| [DellOS10] | Manuel de référence OS10 (OS10.5), pour les commutateurs S4128F-ON | Les commandes, des résumés sur les concepts |
| [DellVLTArch] | Dell Corporation - Architecture VLT | Livre blanc sur l'architecture VLT |
| [PARefPA820] | Référence documentaire des routeurs PA 820: | Documentation en ligne sur le site PaloAlto |

# 2.2 Architecture

## 2.2.1 Vue générale: équipements et niveaux fonctionnels

Les équipements sont dans plusieurs catégories, correspondant à des niveaux fonctionnels de la plateforme, comme représentée dans la figure suivante:



Le tableau suivant liste les principales caractéristiques des équipements

| Equipement | Nb - Modèle | Caractéristiques |
|---|---|---|
| Routeurs<br>RT1, RT2 | 2 X PaloAlto PA 820 | Configuration des équipements réseau |
| Commutateurs<br>SW2, SW3 | 2 X Dell S4128F-ON | 32 ports SFP, commutation 10GbE et 1GbE,<br>support trunking via 2 ports QFP 28 à 100G<br>OS: OS10 Service Fabric 10.5 (Enterprise Edition) |
| Noeuds proxy<br>NP2, NP3 | Dell PowerEdge 440 | 2 CPU, 49Gb RAM,<br>Disque: 2X 300 Gb 15krpm (RAID1), 4 X 485 Gb SSD<br>OS : Windows Server 2019 Standard |
| Noeuds de virtualisation<br>NV3 .. NV6 | Dell PowerEdge 640 | 2 CPU, 192 Gb RAM,<br>Disque = 2X 300 Gb 15krpm, 5 X 485 Gb SSD<br>OS: Windows Server 2019 Datacenter |
| Noeuds de stockage<br>NAS 2, 3 | 2 X Dell PowerVault 3420 | 1 CPU, 8gb RAM<br>Disque:<br>système: 400Gb utiles, type RAID1s/disques 15krpm<br>stockage: 7 To utiles , type RAID5 s/disques 7500 rpm<br>OS: Windows Server Storage 2016 |

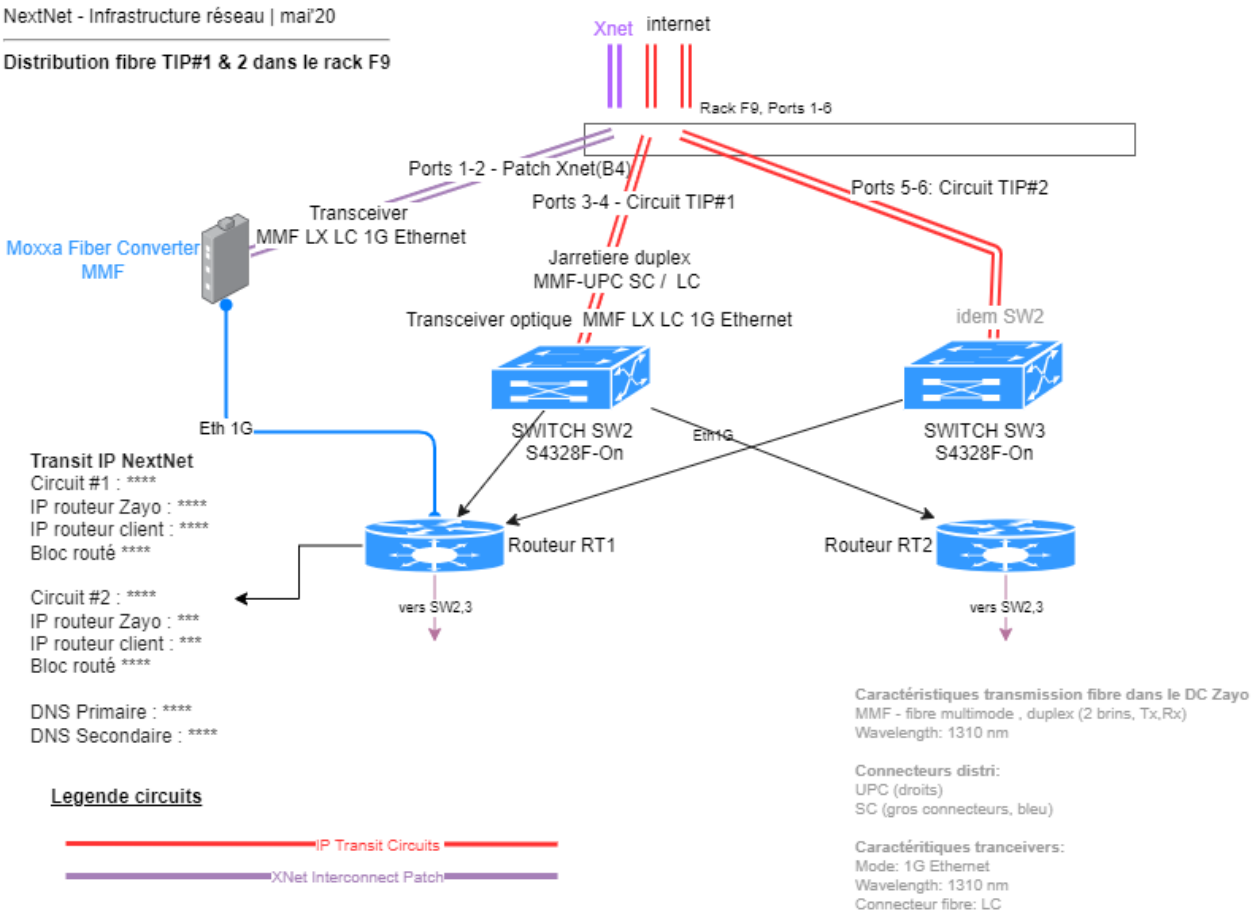
### 2.2.2 La haute disponibilité

L'architecture NextNet vise la haute disponibilité de la plateforme, adressée à travers des choix de configuration au niveaux suivants:

- La configuration et l'utilisation des circuits IP
- Le routage
- La configuration des interfaces réseau
- La commutation

## 2.2.3 La connexion des circuits IP TIP#1 et TIP#2 sur NextNet

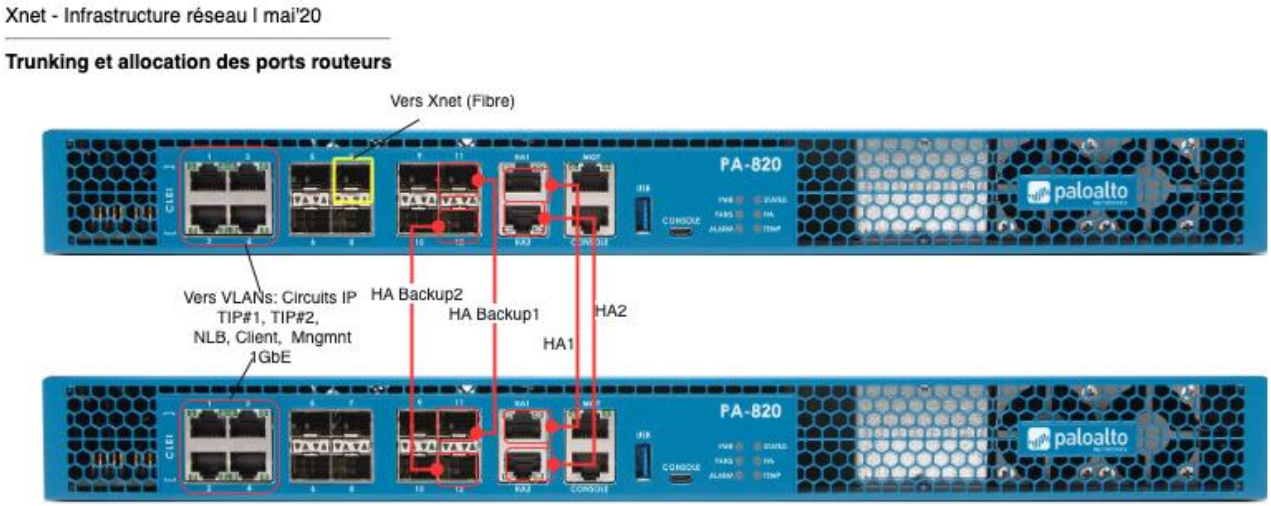
Les circuits ip proposés par le datacenter sont mis en œuvre comme schématisé ci-dessous (pour des raisons de sécurité les adresse IP ne sont pas mentionnées dans ce document):




## 2.2.4 Le routage HA (et ses limites)

Les deux routeurs sont configurés en mode HA, sur le principe actif (RT1)/ passif (RT2). C'est une fonctionnalité native des routeurs PA 820.

Cette configuration nécessite une utilisation parfaitement symétrique des ports des deux routeurs, afin que la panne de l'un ou l'autre des deux routeurs ne bloque pas le fonctionnement de la plateforme.




Les limites de la configuration actuelle sont:

- Chaque circuit de transit ip internet (TIP#1, TIP#2) étant pris en charge par un commutateur seulement, une panne de l'un des commutateurs SW2 ou SW3 rend indisponible le circuit de transit IP correspondant. Les mécanismes de HA DNS doivent assurer le mode dégradé de la plateforme dans ce cas.
- Le patch fibre intraDC avec la plateforme Xnet est prise en charge directement par un routeur et non pas par un commutateur, car le patch n'est pas dual (ie 2 circuits fibre entre les deux racks NextNext et XNet). En cas de panne du routeur actif, la communication NextNet-Xnet est rompue.

## 2.2.5 La configuration des NIC et des ports réseau des serveurs

### *2.2.5.1 Caractéristiques*

Les NIC (cartes réseau) des 3 types de serveurs (NP, NV, NAS) sont dotés de 2 ou 4 ports, ayant la possibilité de communiquer en 1GbE et 10GbE par paire de deux ports.

En règle générale, les communications ont lieu en 1GbE, en sachant que les débits 1GbE sont des débits maximums autorisés au niveau deTIP#1&2 proposés par le datacenter, et des routeurs également.

Le débit 10GbE est réservé aux VLAN  S2D (isolé) pour les serveurs NV, afin de respecter les préconisations pour les architectures hyperconvergentes nécessitant de débits importants pour la réplication des espaces de stockage S2D.

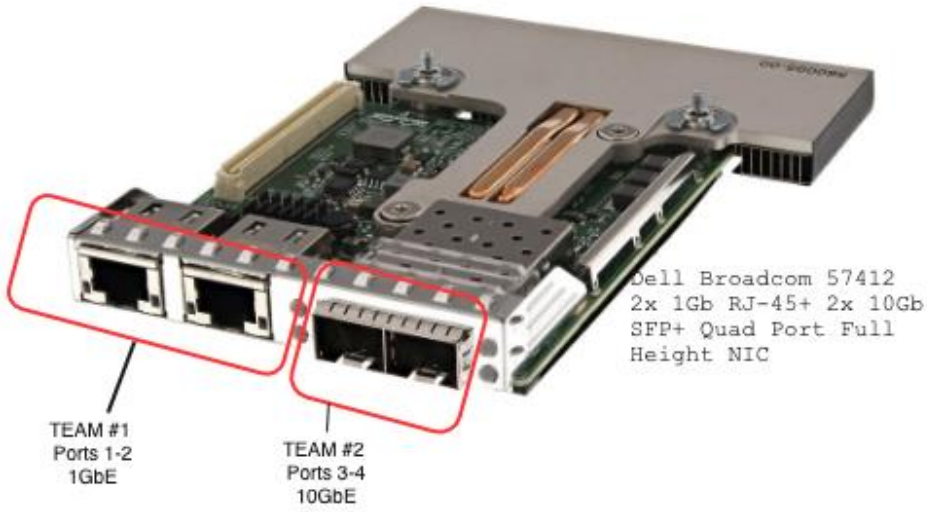
**Observation** : La haute disponibilité ne repose pas sur la présence de NIC multiples sur les serveurs, mais sur la multiplicité des équipements à chaque niveau.

### *2.2.5.2 L'agrégation (teaming) des ports réseau*

Chaque serveur (proxy NP, virtualisation NV, stockage NAS) dispose de minimum quatre ports réseau, portés par un ou deux NIC (cartes réseau).

NextNet - Infrastructure réseau I mai'20

**Allocation des ports réseau pour les serveurs NV**




Les ports sont systématiquement agrégés en paires (teams) de 2 ports, ayant la même bande passante, et chaque port est connecté à un commutateur différent.
Exemple de configuration pour les nœuds NV (virtualisation):

- L'agrégation des ports 1&2 est configurée en 1G et utilisée pour le subnet Client Access & Cluster Private
- L'agrégation des ports 3&4 est configurée en 10G et utilisée pour le subnet Cluster Storage

Ceci permet au réseau de rester opérationnel en cas de perte de 1 commutateur et implicitement la perte d'un circuit IP.

***Limites et facteurs de mitigation***

Il faut observer que le teaming n'adresse pas la perte d'un port réseau (ce qui est plutôt rare) et il n'adresse pas la perte de la carte réseau si les deux ports sont sur une même carte. La perte d'une carte réseau entraîne la perte du serveur. Ceci est mitigé par le fait que la perte de 1 voire 2 serveurs par niveau reste acceptable, voir l'infrastructure applicative NextNet.

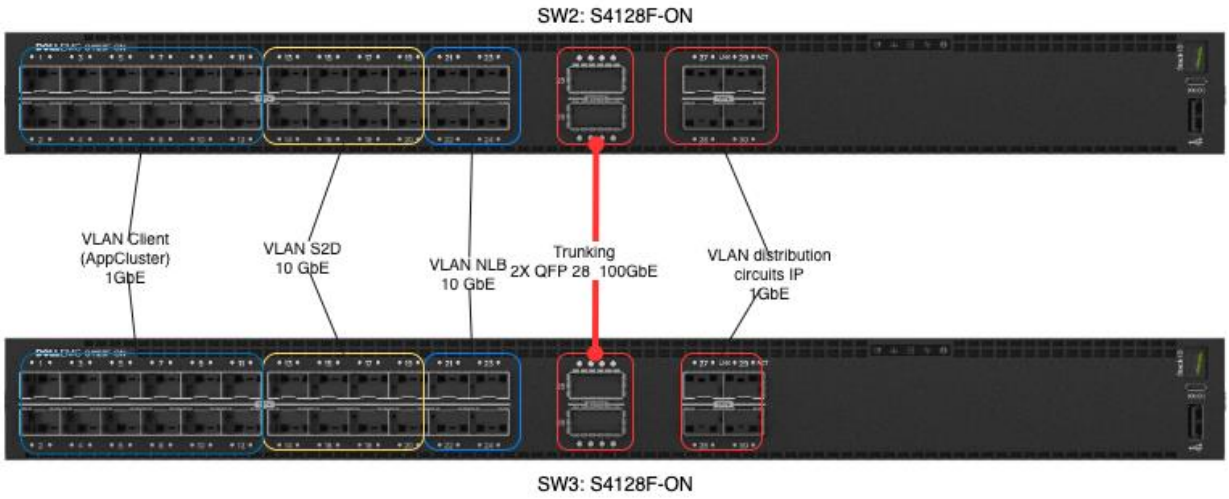
## 2.2.6 Technologie de commutation : le trunking VLT

La commutation est assurée par deux commutateurs (SW2 et SW3 ), de type S4128-F ON, proposant 28 ports SFP+ fonctionnant en 10GbE ou 1GbE et deux ports QFP28 utilisant - dans notre cas - un débit de 100GbE.
Les commutateurs sont agrégés (trunking) à travers les ports QFP28 la technologie d'agrégation VLT (Virtual Link Trunking) proposée par Dell et **l'OS 10 Service Fabric**.

L'agrégation des deux commutateurs permet un fonctionnement HA actif / actif et assure également un équilibrage de charge des flux réseau.
La figure suivante présente l'allocation (symétrique) des ports des deux commutateurs. L'allocation effective des ports est maintenue dans tous les cas.
Observation: L'équilibrage de charge n'est pas forcément une préoccupation forte dans la configuration de la commutation, c'est la haute disponibilité et le failover en cas de panne de l'un des commutateurs qui prime.

Xnet - Infrastructure réseau I mai'20

**Trunking et symétrie VLANs**



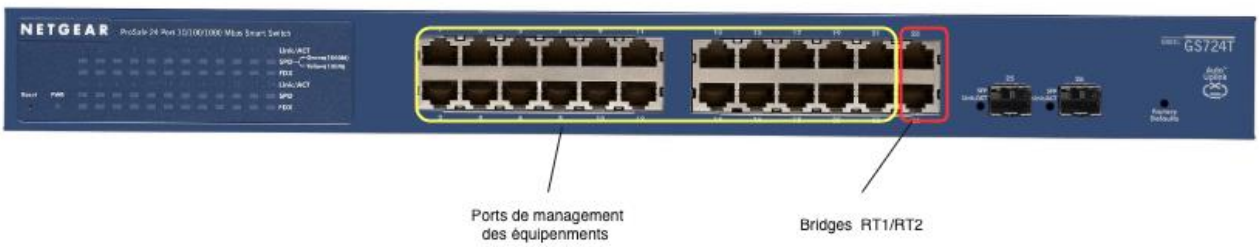

## 2.3 Sous réseau de management

L'ensemble des ports de management des équipements sont commutés par un commutateur low cost de type Netgear GS724T. Sur le schéma suivant on distingue la connexion au bridge de routage sur les routeurs, et aux ports de management des différents serveurs.

Le port de management du commutateur de management lui-même ne bénéficie pas d'un port physique (comme les autres équipements) mais d'une adresse IP.

Xnet - Infrastructure réseau I mai'20

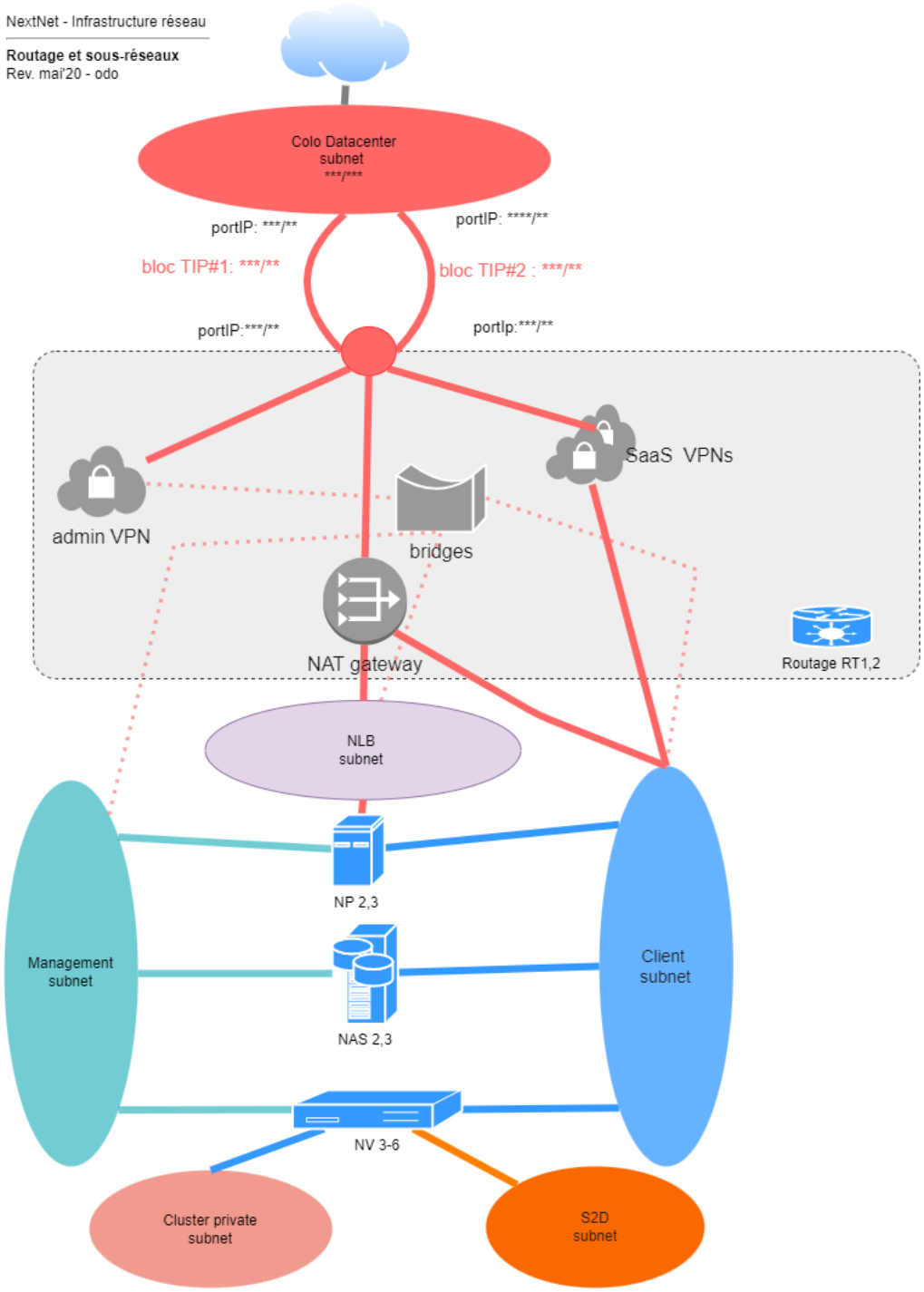
**Allocation des ports de management**



### 2.3.1 Sous réseaux et serveurs

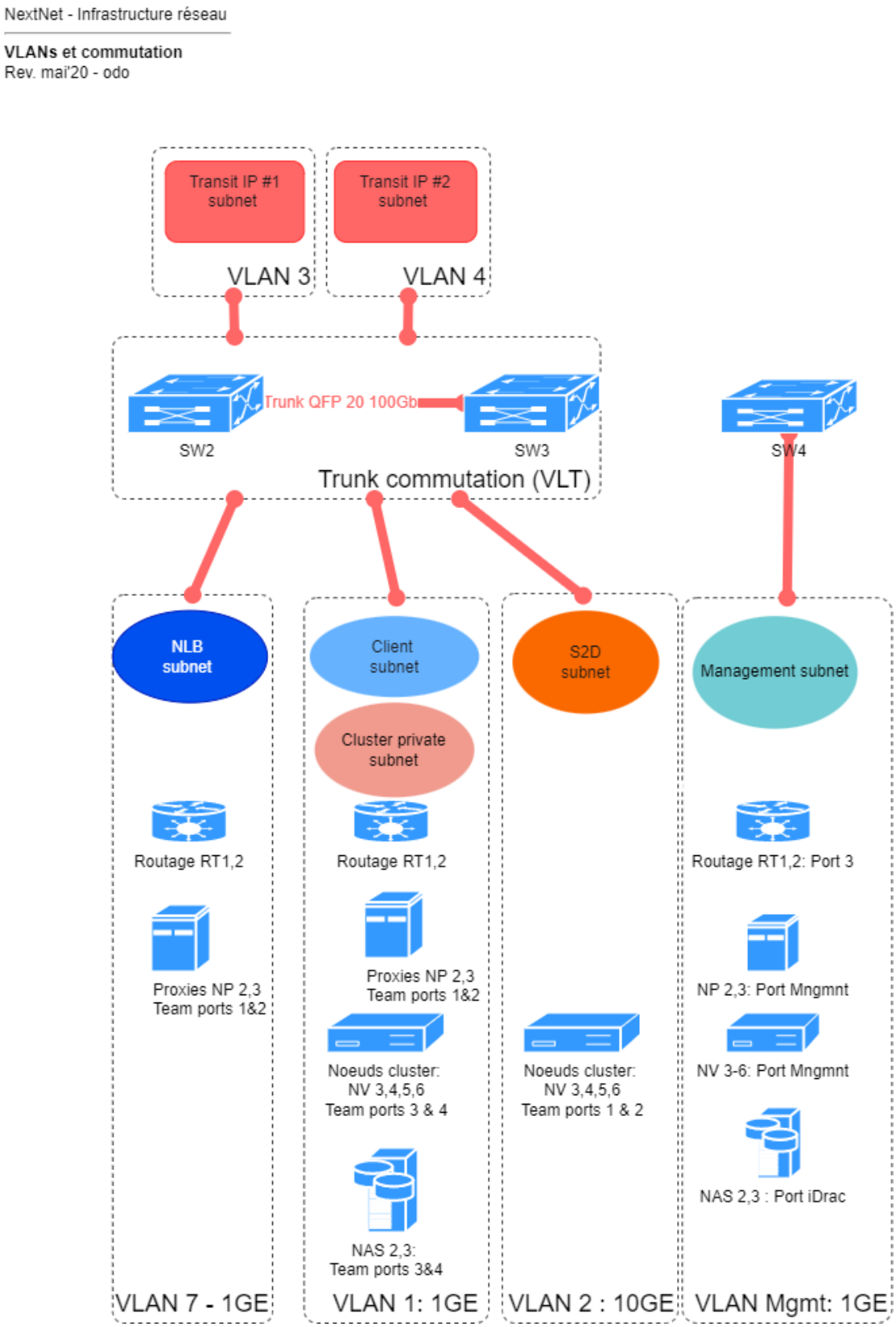
La figure suivante met en évidence les sous-réseaux utilisés, afin de séparer les flux réseau logiques et les équipements qui les utilisent.

On observe les réseaux propres aux serveurs de virtualisation, qui ne sont pas reliés à des ponts réseau au niveau routage: le réseau S2D et le réseau privé du clustering.

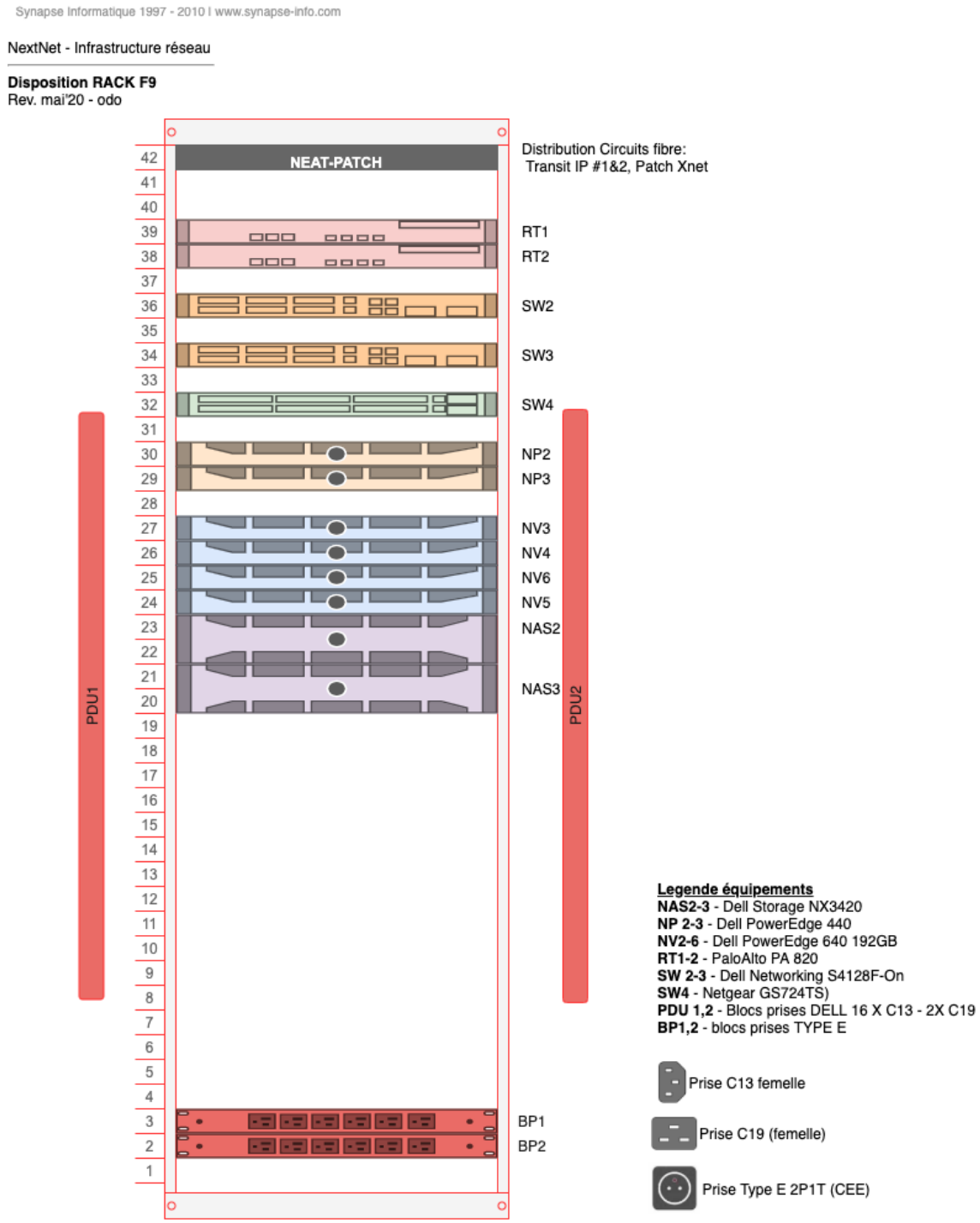


## 2.3.2 La projection sous réseaux -> VLANs

Le domaine de commutation global est divisé en plusieurs VLANs, comme illustré dans le schéma suivant:



## 2.3.3 Le montage des équipements en rack

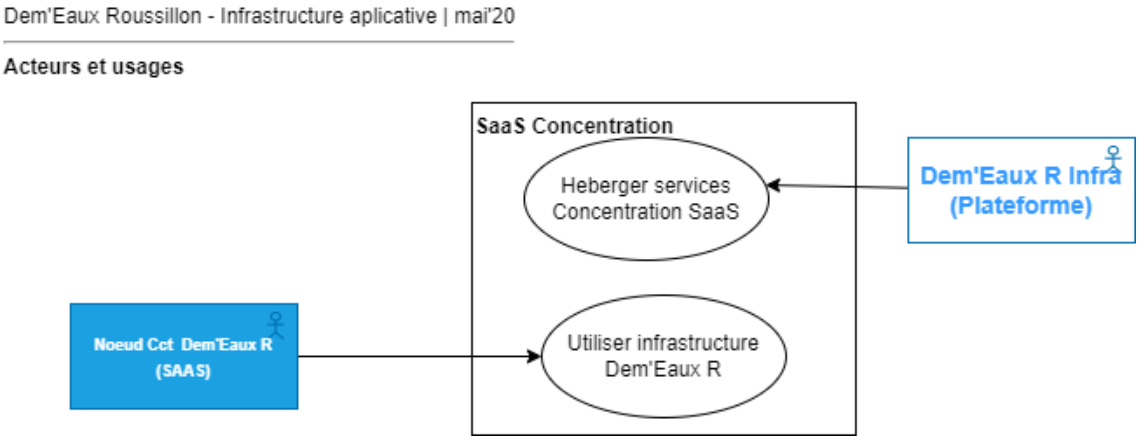


# 3 INFRASTRUCTURE APPLICATIVE

## 3.1 La vue fonctionnelle

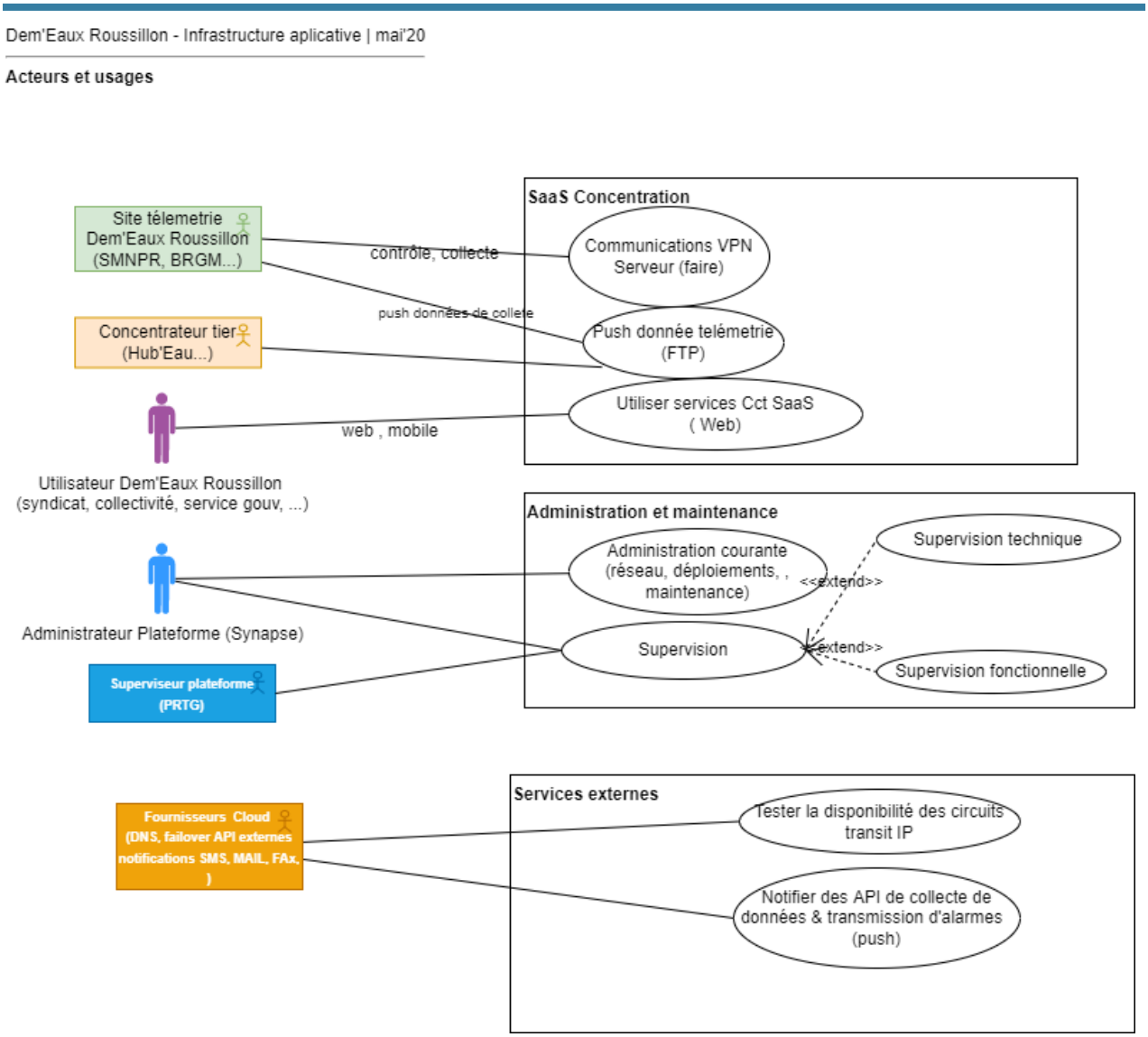
### 3.1.1 Usages et acteurs de la plateforme

La plateforme NextNet - Dem'Eaux Roussillon est une plateforme de production, pour les services de concentration et de valorisation des données, utilisés par les partenaires (producteurs de données, collectivités, organismes de recherche) et les usagers de la plateforme.




Les principaux cas d'utilisation de la plateforme sont:

- L'exécution des services SaaS de concentration de données, incluant la collecte des données, les échanges, les notifications d'alerte
- L'administration et la maintenance de la plateforme
- Les interactions avec les systèmes tiers (et cloud) des fournisseurs internet / télécom, pour la réalisation de certains services fonctionnels (collecte, notification d'alerte) ou le monitoring de la disponibilité (fournisseur DNS)




# 4 ARCHITECTURE GLOBALE

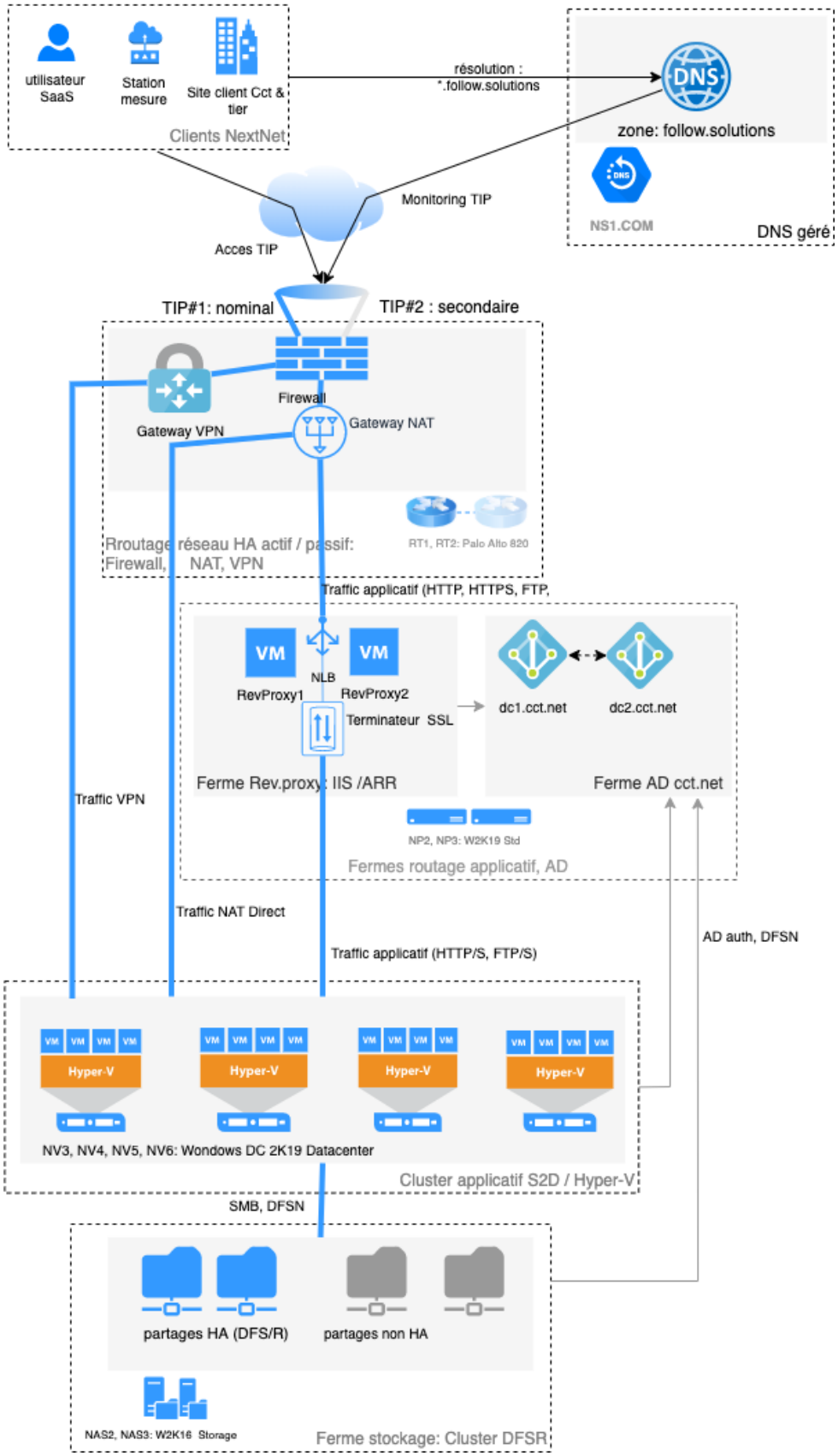
Elle est décrite en plusieurs niveaux

- L'accès à la plateforme (routage externe)
- Le routage applicatif
- Le Cluster applicatif
- Le Stockage de contenu

Ces niveaux sont représentés sur le schéma suivant, indiquant pour chaque niveau le support d'exécution matériel (machines, équipements réseau), les technologies et mécanismes de haute disponibilité utilisés, les principaux flux de communication entre les niveaux.

NextNet - Infrastructure aplicative I mai'20

**Vue d'ensemble**



# 4.1 L'accès externe à La plateforme

## 4.1.1 Les points d'accès publics et les transits IP en DataCenter

Les services Dem'Eaux sont accessibles sur internet à travers deux plages d'adresses IP /24 (8 adresses):

- une plage d'adresse primaire – le transit ip n°1 ( tip#1 ),
- une plage d'adresse secondaire - le transit ip n°2 (tip#2)

Ces plages correspondent au niveau réseau à deux circuits physiques de transit ip (tip), matérialisés par des connexions physiques distinctes de fibre optique vers NextNet.

***Pourquoi deux TIPs pour la colo NextNet?***

L'utilisation de deux TIPs est en relation avec la défaillance des équipements réseau qui reçoivent les connexions TIPs et de la responsabilité des pannes.

En effet, avec les solutions et connaissances que nous avons actuellement, si un commutateur est en panne, le TIP n'est plus exploitable sans intervention humaine, ce qui n'est pas acceptable pour des services de concentration temps réel.

Le cas de la panne du TIP à cause d'une défaillance extérieure à l'installation NextNet-DemEaux - et hors champs d'action Synapse, comme une panne d'équipement dans la MMR du datacenter par exemple ou une panne de l'alimentation du datacenter - n'est pas adressée par NextNet-DemEaux actuellement. L'objectif étant de rapprocher le niveau de résilience de la location NextNet-DemEaux le plus possible de celui du datacenter de colocation.

## 4.1.2 DNS failover et politique de noms des services applicatifs

Les points de terminaison de service NextNet-Demeaux sont accessibles à travers des noms du domaine comme **follow.solutions ou followtest.solutions** pour les instances de test ou de préproduction comme "<**demeauxroussillon**>**.follow.solutions" ou <demeauxthau>.follow.solution**.

La politique HA pour les accès aux services publics NextNet est la suivante:

- Les points de terminaison (adresses) des services HA sont traités systématiquement via le tip#1, dit nominal, ou primaire
- En cas de défaillance du tip#1, le service est accédé via le tip#2, dit secondaire

Le fournisseur DNS de **follow.solutions** est géré **ns1.com,** fournisseur mettant un mécanisme de basculement depuis la perspective d'un client de la plateforme.

Ce mécanisme fonctionne comme ceci:

- L'enregistrement DNS de résolution **(*.follow.solutions**) dispose de deux réponses possibles: une adresse sur tip#1:n et l'adresse homonyme tip#2:n. Le TTL du record DNS est fixé à une durée suffisamment basse (300 secondes = 5 minutes), correspondant au temps acceptable d'un basculement depuis la perspective du client final en cas de défaillance du tip#1.
- Le service DNS supervise en permanence l'état de santé du tip#1 avec des requêtes (ping, avec des limites sur les temps de réponse).
- Lors d'une demande de résolution, le service DNS répond avec l'adresse du tip#1 si la santé de tip#1 est bonne, sinon avec l'adresse du tip#2.

## 4.2 Le routage applicatif

Le routage applicatif est réalisé avec des solutions de routage et équilibrage de charge qui lui sont propres (NLB, ARR), et il est basé sur le routage et la commutation au niveau du réseau, décrit dans **Erreur ! Source du renvoi introuvable.**.

### 4.2.1 Flux applicatifs entrants et leur traitement

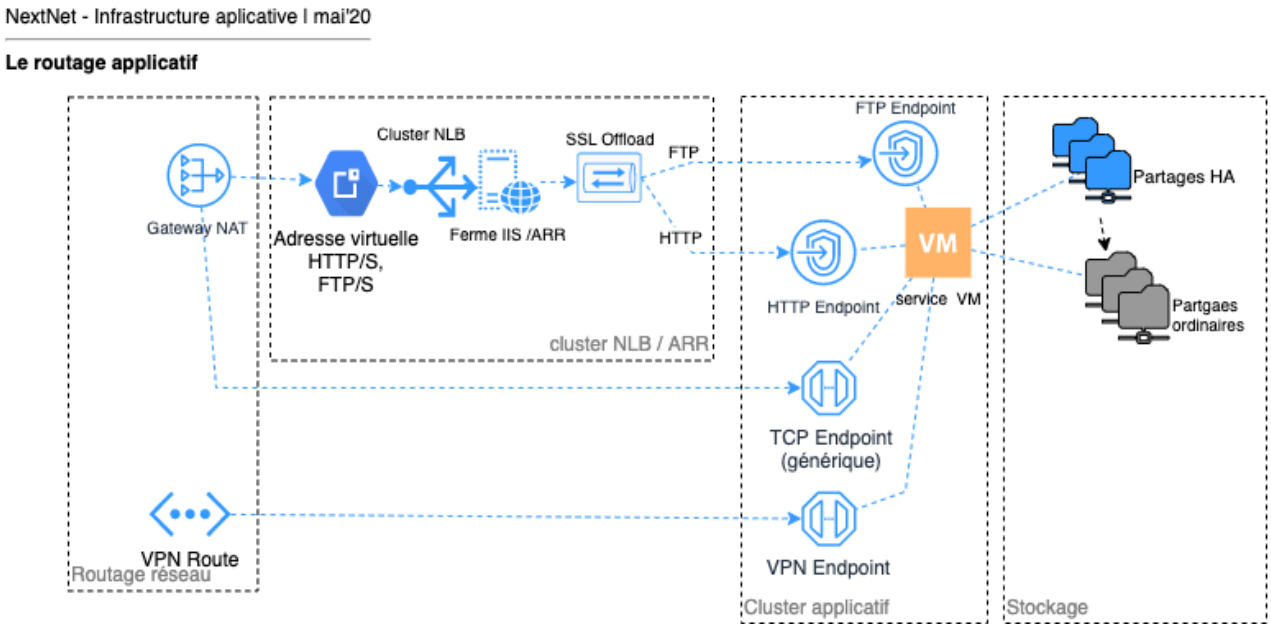
Nous avons plusieurs types de flux

- Le traffic Web externe (http, https)
- Le trafic FTP/S
- Les flux VPN, en deux catégories :
  - Les VPN permanents de site à site (IPSEC) – utilisés pour des échanges permanents de données
  - Les VPN client / serveur, exposés par un serveur NextNet pour assurer de la télémétrie en général ou des accès administratifs
- Les backdoors (temporaires), comme celui avec Xnet, permettant l'accès direct entre les deux plateformes
- Autres types de trafic spécifiques, de type TCP / UDP en général

### 4.2.2 Le cluster NLB

Le cluster NLB assure en premier lieu **la disponibilité** en cas de panne des serveurs frontaux et en deuxième lieu l'équilibrage de charge entre les deux serveurs. Les flux concernés sont :

- HTTP/S (port 80/443) vers les 2 instances ARR (reverse proxy), puis vers les endpoints applicatifs concernés
- FTP/S (port 21 / 990) vers les instances applicatives dédiés au service FTP




### 4.2.3 Le Périmètre NLB

Les protocoles VPN et les autres protocoles TCP (VPN, TCP ordinaires, UDP, …) ne sont pas pris en charge par NLB dans la version actuelle. Ces flux sont routés directement vers des endpoints applicatifs

### 4.2.4 Configuration réseau pour NLB

***Le multcast IGMP***

L'option "multicast IGMP" pour la gestion des communications NLB est la méthode de gestion réseau la moins impactante d'un point de vue trafic, mais elle nécessite que les commutateurs supportent IGMP v2. Ainsi, IGMP v2

doit être activée au niveau des commutateurs réseau, par l'activation du "IGMP Snooping" (le fouinage IP) par le commutateur afin d'optimiser le multicast. Voir **Erreur ! Source du renvoi introuvable.** pour la configuration réseau.

***Implémentation des nœuds NLB avec des VMs (Hyper-V)***

Si le NLB est pris en charge au niveau des VMs (versus machines), certains points seront également vérifiés:
https://techcommunity.microsoft.com/t5/failover-clustering/deploying-network-load-balancing-nlb-and-virtual-machines-on/ba-p/371631#

https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/techpaper/implmenting_ms_network_load_balancing.pdf

### 4.2.5 Le trafic HTTP/S: SSL offload, ARR

Le flux Http/S est généré par les clients web interactifs des services SaaS ou par des backend tiers intervenant dans la concentration comme:

- Les sites de mesure
- La notification des backend des fournisseurs de service avec des clients HTTPS pour:
  - La collecte de données (sites de collecte primaire, comme les backend IoT)
  - Les notifications de transmission d'alarme
  - etc

Le routage HTTPS est réalisé par une ferme de deux serveurs Windows IIS / ARR (avec le module URL rewrite). Les deux serveurs ont des interfaces réseau sur le sous-réseau de routage et le sous-réseau du cluster applicatif.
L'ensemble de ces flux sont (doivent être) sécurisés en utilisant HTTPS.
La ferme ARR assure également un **point de terminaison SSL** (SSL offloading), i.e. les flux redirigés vers et les points de terminaison applicatifs exposés par les hôtes de services du cluster applicatif.

### 4.2.6 Les flux FTP/S

Les flux FTP/S sont routés à travers la ferme ARR, vers des VM applicatives dédiés au service FTP.

## 4.3 Le cluster applicatif: S2D – Hyper-V

Le cluster applicatif est basé sur un cluster S2D / Hyper-V, assurant une triple redondance (3 way mirroring) des VM.

Le cluster assure une continuité de service des VMs lors de la panne des 2 / n noeuds physiques, avec un minimum de 4 nœuds physiques minimum.

Les applications déployées dans le cluster sont:

- des noeuds de concentration SaaS, incluant toutes ou une parties des modules de concentration: web, collecte, alarme, échange, me
- des frontaux de télémétrie (FTP, http)
- des nœuds partiels de concentration,
- etc

Une instance d'application s'exécute sur une VM du cluster.

Une VM peut exécuter une ou plusieurs instances applicatives.

## 4.4 Le stockage : cluster DFS et ressources de stockage

### 4.4.1 Technologie

Le niveau stockage est assuré par une ferme d'Appliance NAS basés sur Windows Server Storage 2016, fournissant des ressources de stockage accessibles en SMB à travers DFSN (DFS Namespace).

Les noms DFSN sont associés à des espaces de stockage faisant l'objet d'une réplication DFSR.

**Remarque**: la réplication DFSR est une technologie de réplication de niveau applicatif – qui à la différence du mirroring – assure la réplication des fichiers. La réplication a lieu une fois lorsque que les fichiers ne sont plus accédés en écriture. La conséquence de ce comportement est que la technologie n'est pas adaptée pour la sécurisation du contenu des fichiers restant ouverts sur des longues périodes.

### 4.4.2 Ressources de stockage

Les ressources exposées par la ferme NAS sont dans les catégories suivantes:

- Ressources critiques – implémentée sur des espaces DFSR
  - Espaces de stockage pour les fichiers de télémétrie, utilisés par des front ends **http** ou **ftp**.
  - Sauvegardes des données
- Ressources non critiques – implémentées en tant que partages ordinaires par l'un des serveurs NAS
  - Espaces de travail temporaire pour les applications
  - Media / images système permettant l'administration de la plateforme
  - Sauvegardes temporaires

Voir le plan de déploiement applicatif pour le dimensionnement des ressources de stockage

## 4.5 Les services applicatifs de concentration – Typologie déploiement, dimensionnement

### 4.5.1 L'architecture mono tenant de la solution de concentration

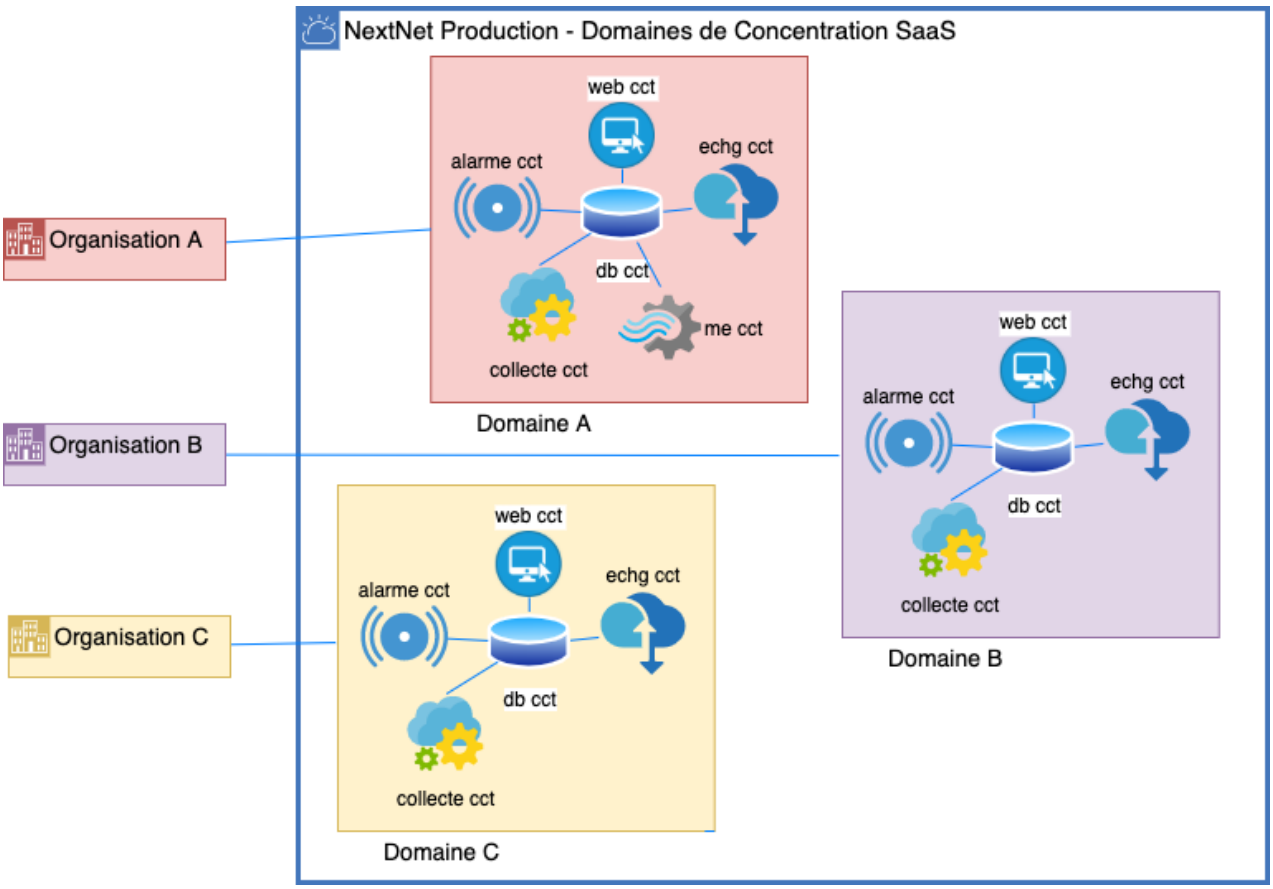
Un **domaine de concentration est défini comme étant l'ensemble** des fonctionnalités et les données associées à une organisation cliente des services de concentration NextNet. Le domaine de concentration est pris en charge par une instance applicative de concentration comprenant plusieurs modules.

La solution de concentration présente - dans la perspective d'un déploiement SaaS et de la mutualisation Dem'Eaux Thau et Dem'Eaux Roussillon - une architecture mono tenant, permettant de traiter un seul domaine de concentration avec une instance donnée de module comme la base de données ou un des services de concentration.


Plus précisément :

- Une instance de base de données – correspond à un seul domaine de concentration (Thau ou Roussillon) ; une instance de moteur de données peut cependant gérer des instances multiples de bases de données de concentration (par exemple base de test ou de préproduction).
- Une instance d'un services de concentration (collecte, alarme, échange , etc.) traite un seul domaine de concentration également; le traitement de multiples domaines de concentration nécessite des instances multiples de ces services.

Cette situation est représentée dans le schéma suivant



Même si l'architecture mono-tenant ne favorise pas la mutualisation à grande échelle, elle présente l'avantage d'une gestion flexible des versions des installations et de limiter les impacts des éventuels incidents de fonctionnement de la plateforme.

### 4.5.2 Types et gabarits de VM

Les OS invités proposés par le cluster applicatif sont dans les catégories:

- **Windows Server** - Windows 2019 Datacenter
- **Linux** - Ubuntu

Voir https://docs.microsoft.com/en-us/windows-server/virtualization/hyper-v/supported-windows-guest-operating-systems-for-hyper-v-on-windows pour des précisions.

Afin de standardiser l'allocation des ressources comme le processeur, la mémoire ou l'espace de stockage, plusieurs gabarits de VM sont définis, avec des caractéristiques amenées à être rectifiés selon le niveau de performance observé en production. Cette approche permet de redimensionner une VM si nécessaire en fonction de l'évolution du besoin : volume de données croissant, augmentation de la fréquentation et du nombre d'utilisateurs…

Le gabarits définis sont

- **Basic (BVM )** – gabarit réservé aux déploiements dans un but de test ou expérimental
- **Standard, Medium, Large** (SVM , MVM, LVM) – gabarits de production, pour exécuter une ou plusieurs installations de concentration selon leur profil fonctionnel (S, M, H) et de déploiement (SC ou MC, voir ci-dessous).

Les caractéristiques approximatives des gabarits VM sont listées dans le tableau suivant. Voir **Erreur ! Source du renvoi introuvable.** pour la référence des gabarits.

| Gabarit VM | Processeur | RAM | Stockage (unités 20GB) |
|---|---|---|---|
| Basic | 1 VCPU | 2Gb | 1..2 |
| Standard | 2 VCPU | 6Gb | 2..6 |
| Medium | 4 VCPU | 12Gb | 4..10 |
| Large | 6 VCPU | 16Gb | 6..12 |

Les caractéristiques suivantes:

- **interface réseau** (1/10GbE)
- **stockage local SSD**
- haute disponibilité des conteneurs VM (failover automatique des VM sur le cluster S2D / Hyper-V),

sont communes à toutes les VM fonctionnant sur le cluster applicatif.

Les plages pour l'espace de stockage sont déterminées par des facteurs de périmètre fonctionnel (données scalaires, image), les contrats de service (l'antériorité des données en ligne), etc.

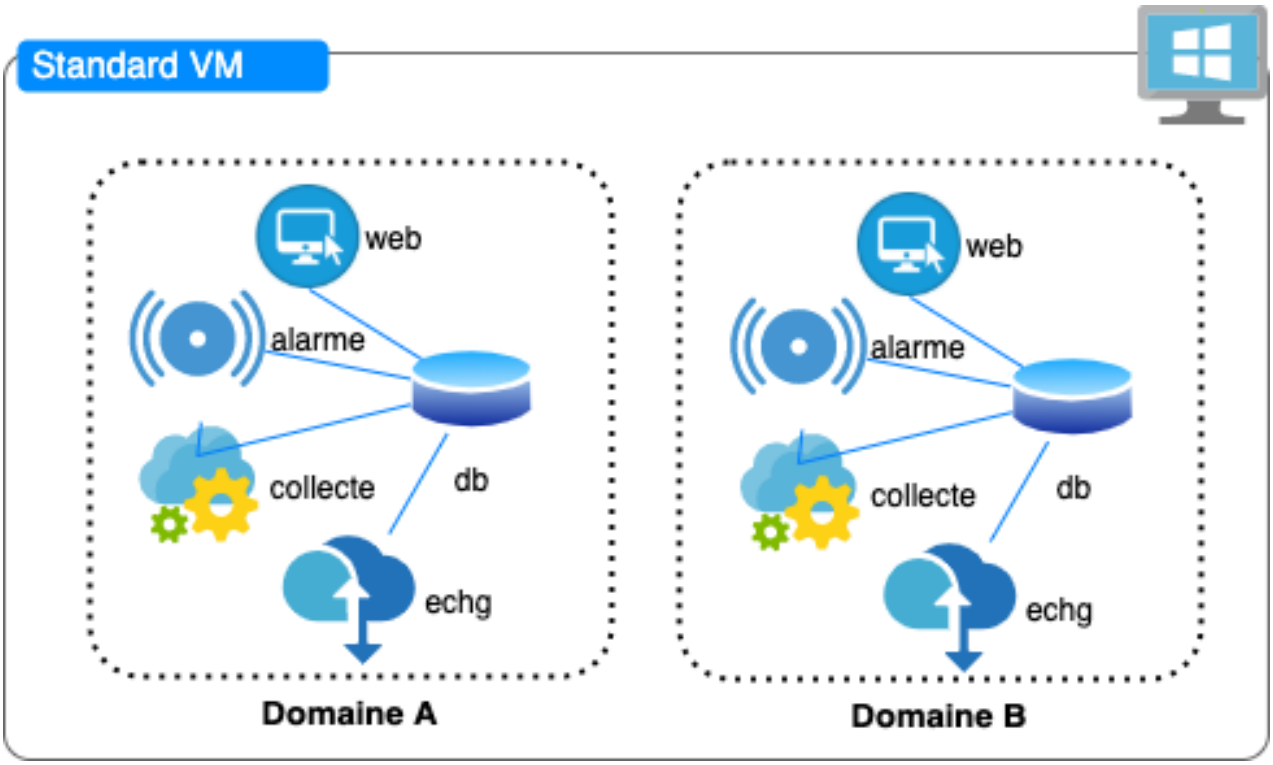
Il est néanmoins indiqué ici dans des incréments de 20 Gb, dans la mesure où on observe une corrélation dans le fonctionnement de la concentration entre les volumes de stockage et la consommation des ressources CPU et RAM.

## 4.5.3 Modèles de déploiement des domaines de concentration

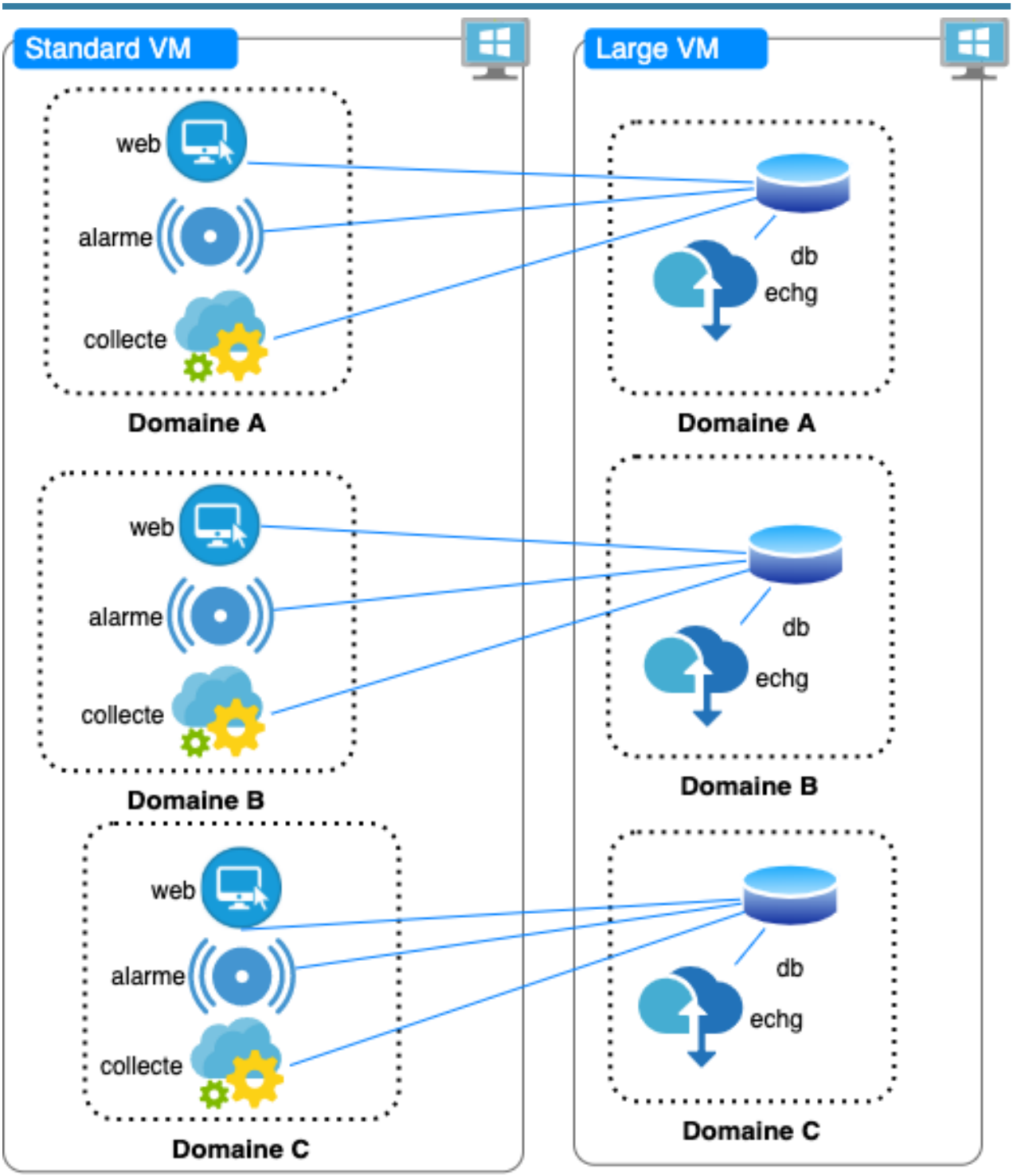
### *4.5.3.1 Modèles de déploiement: SC , MC*

Deux modèles de déploiement sont proposés pour les installations de concentration Dem'Eaux :

- Le modèle SC (single container) - l'ensemble des modules de l'installation sont déployés sur une même VM, qui est mutualisé. C'est actuellement le cas pour les sites Dem'eaux Thau et Dem'eaux Roussillon.

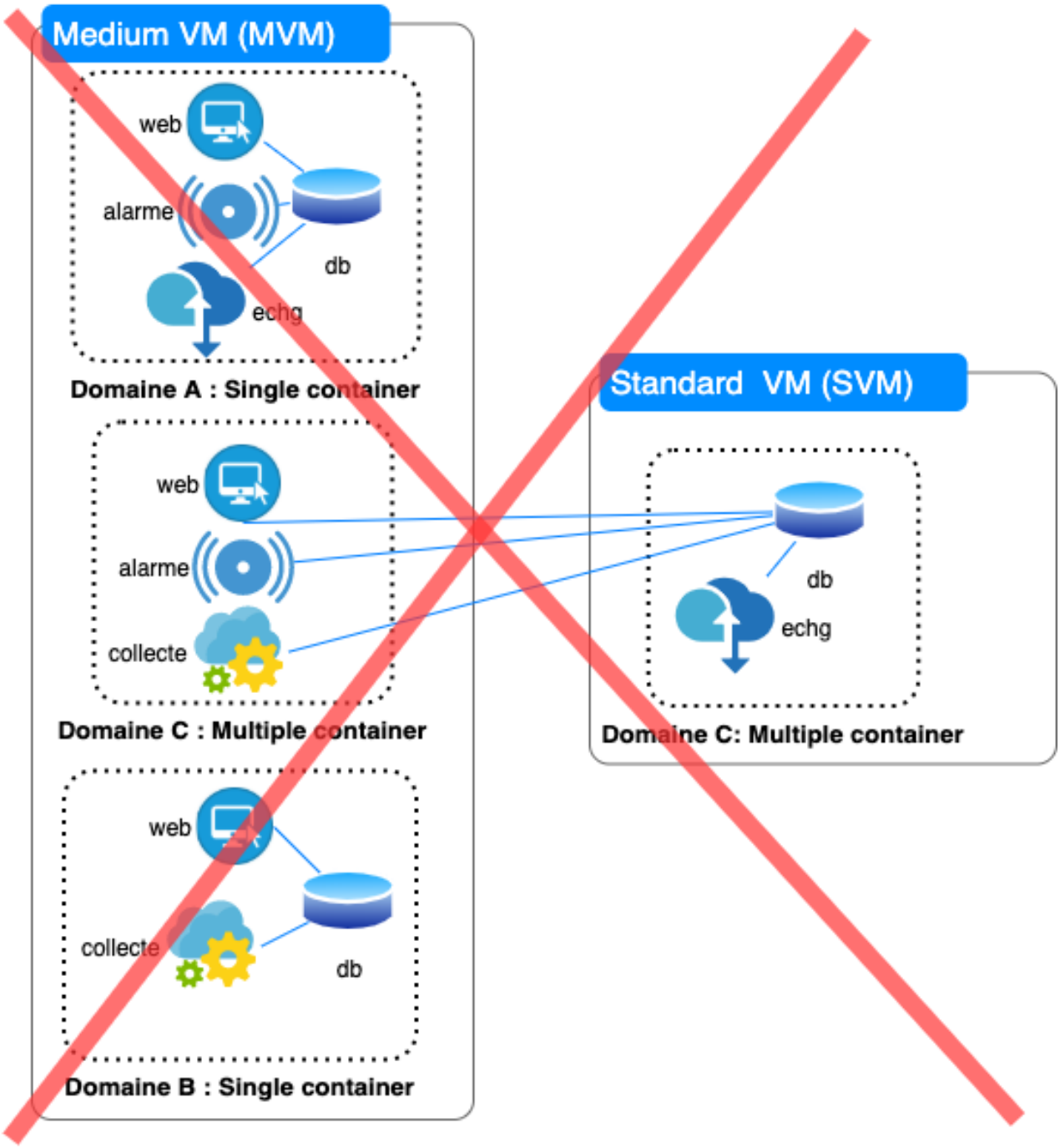



- **Le modèle MC (Multi container)** – les modules sont déployés sur deux ou plusieurs VM, afin de déployer séparément la base de données notamment. Les installations de concentration Dem'eaux pourront évoluer vers ce modèle en fonction des nouveaux besoins identifiés.




### 4.5.4 Mutualisation des VMs

Une VM donnée peut être mutualisée pour des installations adoptant le même modèle de déploiement (SC ou MC); autrement dit, une VM ne va pas mixer des domaines de concentration déployés sur les deux modèles, à conteneur unique ou multiple. Comme exemple, le déploiement représenté ci-dessous est illégal.




# 5 SERVICES APPLICATIFS TRANSVERSES

Nous trouvons dans cette catégorie des services applicatifs qui sont transverses aux domaines dem'eaux Thau et Roussillon, dont les plus remarquables sont:

- Service FTP / FTPS, assurant le frontal pour la remontée des données de télémétrie
- Les serveurs GeoServer pour la fourniture des couches SIG de plusieurs domaines de concentration

D'autres services en perspective se trouvent dans la même catégorie.

- Des passerelles de télémétrie pour interfacer les backends opérateurs pour des services comme:
  - La transmission des données satellite
  - La transmission des données IoT sur des réseaux comme LoRA ou SIGFOX
  - La gestion des notifications vocales appels vocaux
  - Etc

- Des passerelles ou serveurs mutualisés pour la gestion des transmissions pour les transmissions d'alarmes vocales, SMS, push mobile, etc

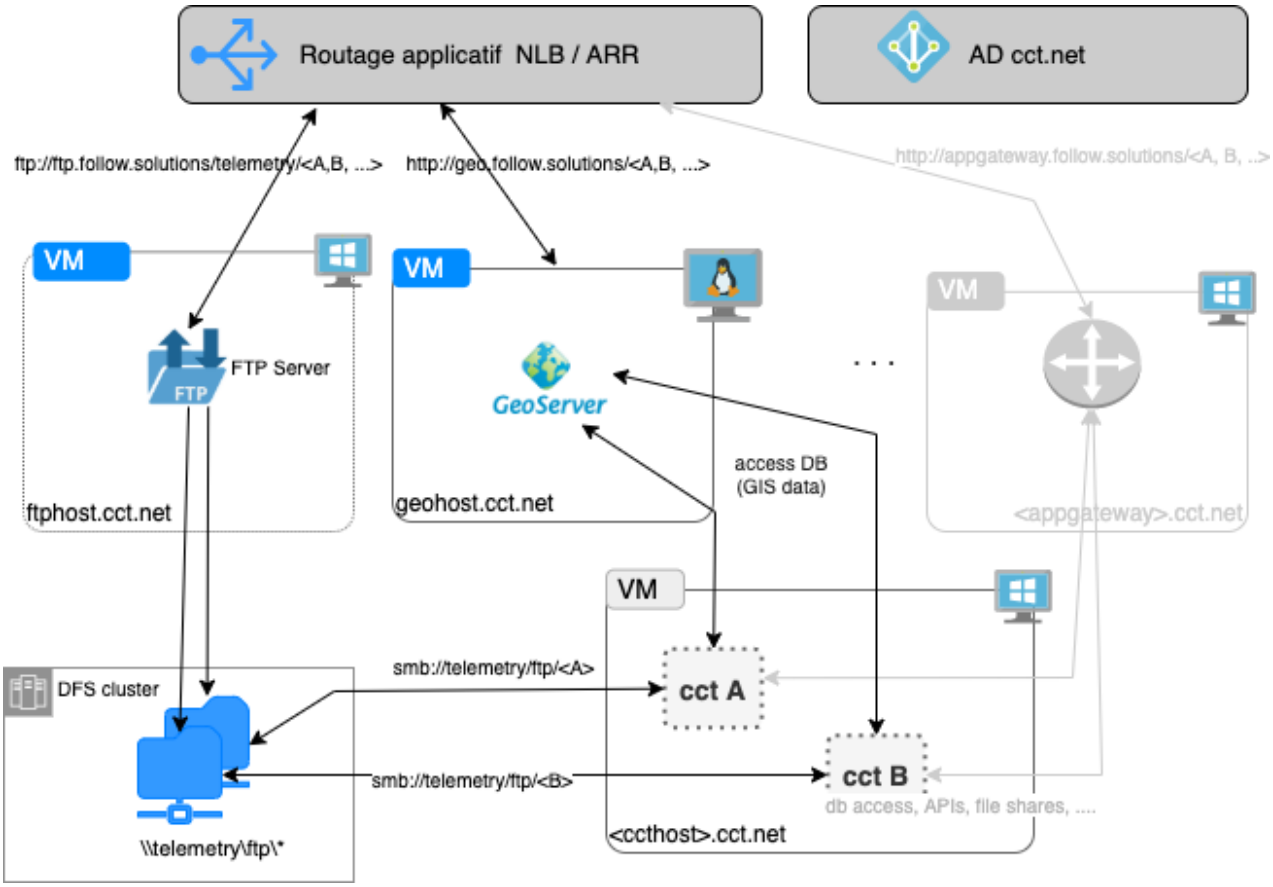
## 5.1 Modèle de déploiement des services transverses

L'isolation des services transverses est au niveau conteneur d'exécution, la VM actuellement.

Les services transverses peuvent être déployés selon le modèle SC (1 service par VM), qui est le modèle nominal. Exceptionnellement, le modèle MC peut être utilisé en sachant que son utilité n'est pas évidente par rapports aux services transverses actuels.

## 5.2 Déploiement des services transverses (ftp, geoserver, ..)

La figure suivante schématise le déploiement des services FTP et Geoserver en mode SC (VM dédiée) et les principaux flux les concernant




### 5.2.1 Le service FTP

Le service FTP est

- Déployé sur une VM dédié W2K19DC (SVM)
- Implémenté sur des sites IIS / FTP avec de **répertoires virtuels, associés à des dossiers dédiés aux domaines de concentration**
- situé derrière le routage applicatif
- utilise l'authentification AD pour les comptes utilisateurs,
- utilise comme espace de stockage des partages DFS

### 5.2.2 Le service de cartographie geoserver

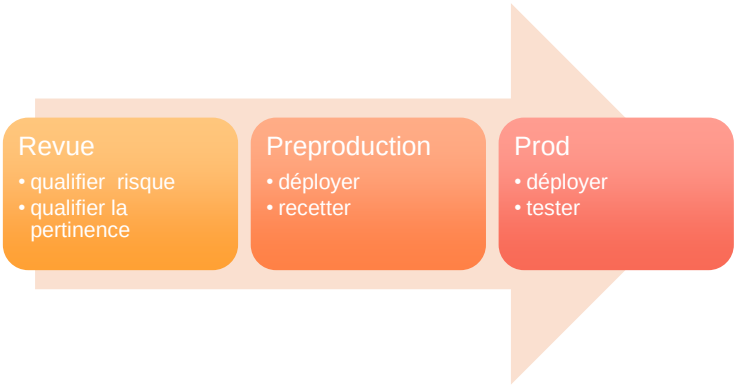
Le service geoserver est:

Déployé sur une VM dédiée utilisant la distribution Ubuntu

- Implémenté sur GeoServer / PostGIS
- situé en aval du routage applicatif
- utilise l'authentification AD pour les comptes utilisateurs
- utilise son espace de stockage local pour le contenu statique
- accède aux instances cct (base de données) concernées pour servir les requetés client relatives aux couches SIG dynamiques.

# 6 POLITIQUE DE MISE A JOUR

Compte tenu des nombreuses technologies mises en œuvre et des risques encourus lors de la mise à jour des systèmes d'exploitation ou des appliances, les mises à jour automatiques des machines physiques et virtuelles de production comporte un risque de dysfonctionnement et d'apparition d'incidents d'exploitation.




Le dispositif de mise à jour repose sur l'application des mises à jour en plusieurs étapes :

- Qualification des mises à jour par
  - une revue du périmètre
  - une revue de la pertinence
  - une revue du risque de dysfonctionnement
- mise à jour sur la plateforme de dev/preprod
- mise à jour sur la plateforme de prod

## 6.1 Périmètre d'application

| Mise à jour | Remarques |
|---|---|
| **Firmware et accessoires des équipements** | |
| Routeur physique | |
| Switches et accessoires | OS pour les switches S 4129F-On<br>SFP<br>Câbles réseau |
| Noueds proxy | Firmware PowerEdge 440<br>Firmware iDRAC |
| Noued de calcul | Firmware PowerEdge 640<br>Firmware iDRAC |
| Appliances NAS | Firmware appliance<br>Firmware iDRAC |
| **Systèmes d'exploitation** | |
| **Windows Server Standard,** | Serveurs physiques NP |

| | |
|---|---|
| **Windows Server Datacenter** | Serveurs physiques NV |
| **Windows Server Storage** | Appliances NAS |
| **Technologies clé** | |
| **Active Directory** | Les VMs AD cct.net |
| **S2D** | Les machines NV du cluster applicatif |
| **DFSR / DFSN** | Les appliances NAS (NASn) |
| **Hyper-V** | Les machines de type NP, NV |
| **Services de routage** | |
| **DNS follow.solutions** | Les VMs AD cct.net |
| **S2D** | Les machines NV du cluster applicatif |
| **DFSR / DFSN** | Les appliances NAS (NASn) |
| **Hyper-V** | Les machines de type NP, NV |
| | |

## 6.2 Les limites de la plateforme de dev/preprod

Les limites suivantes de la plateforme de qualification nécessitent la mise en place de solutions de simulation et/ou d'adaptation du processus nominal

| **Limite** | **Actions et solutions alternatives** |
|---|---|
| Absence routeur physique du même modèle | Revue approfondie, contact fournisseur pour qualification |
| 1 seul switch | Revue approfondie, contact fournisseur pour qualification des dysfonctionnements des fonctionnalités de trunkink |
| 1 seul serveur proxy | Impact:<br>Simulation 2ème serveur proxy avec une VM équivalente |
| 1 seul NAS | Simulation 2ème NAS avec une VM |
| 1 seul domaine DNS de production | Utilisation du domaine DNS de test ( exemple: **followtest.com** ) |
| … | |

# 7 LE MONITORING DE LA PLATEFORME

## 7.1 Le monitoring interne, intra applicatif

Le monitoring interne est aligné sur les indicateurs de monitoring généraux implémentés par les services de concentration.

## 7.2 La supervision externe de la plateforme

Le monitoring externe, réalisé avec PRTG, étend le périmètre de supervision XNet à NextNet.

## 7.3 Le monitoring NextNet avec ELK

L'ingestion des journaux d'activité avec le cluster ELK concerne:

- Les journaux des logiciels de base: **Windows, IIS, SQL Server**
- Les journaux des modules de concentration: **web, collecte, alarme, échange, me**

## 7.4 Le registre d'allocation des ressources applicatives: DE016

Le registre d'allocation de ressources NextNet **Erreur ! Source du renvoi introuvable.** est un document opérationnel permettant de planifier et administrer l'allocation des ressources de la plateforme. Ces ressources sont dans les catégories suivantes:

- Les VM, leur association sur les nœuds d'exécution, leur dimensionnement, leur adressage
- L'allocation des enregistrements DNS
- L'allocation de l'espace de stockage
- etc