MINISTÈRE
DE LA SANTÉ
ET DE LA PRÉVENTION

Liberté
Égalité
Fraternité

Etude d'hébergement de l'entrepôt de données de santé EMC2 dans un environnement souverain

Paris, le 13/12/2023

Délégation ministérielle
au numérique en santé

Rapport d'expertise technique

Rapport relatif à l'étude de faisabilité d'héberger l'entrepôt de données de santé du projet EMC2 sur une plateforme souveraine

Rapporteur : **Héla Ghariani (DNS)**

Avec le concours de :

Emmanuel Clout (DNS)

Vincent Coudrin (DINUM)

Frédéric Law-Dune (ANS)

Maxime Dénès (INRIA)

Sommaire

I. Contexte

a) Contexte du projet EMC2
b) Conclusion du projet de délibération de la CNIL
c) Décision de lancement de la mission EMC2 et objectifs de la mission EMC2

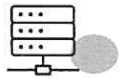
I. A. Contexte du projet EMC2

L'Agence européenne du médicament (EMA) a retenu, fin 2021, la candidature déposée par le Groupement d'intérêt public dénommé « Plateforme des données de santé » (GIP PDS - appelé HDH dans ce document) dans le cadre d'un appel d'offres visant à constituer un **entrepôt de données de santé** multi-centrique afin de permettre la réalisation de recherches, d'études et d'évaluations dans le domaine de la santé, **dénommé « EMC2 »**.

Les données mises à disposition au travers de l'entrepôt EMC2 seront destinées aux établissements de santé, autorités et institutions publiques, associations de patients ou encore les agences européennes comme l'EMA.

Les parties prenantes

- **Pilotage du projet** : Le Health Data Hub (HDH)
- **Fournisseurs de données**
  - Le Centre de Lutte Contre le Cancer Léon Bérard
  - CHRU de Nancy
  - Hospices civils de Lyon
  - Groupe Hospitalier Saint Joseph
  - Plateforme de recherche BPE de l'Université de Bordeaux



Le projet EMC2

- Financement à hauteur de **1,5 M€** sur une période de **6 ans**
- Données de **300 000 à 500 000 patients** / an
- Population témoin avec un ratio de **1 patient pour 3 témoins**
- Appariement des données patients aux **données du SNDS (copie partielle SNDS sur la PDS du HDH)**
- Livraison HDH à EMA : à **mi-Juillet 2024**

Environnement d'hébergement

- Choix initial de Microsoft Azure pour l'hébergement de l'entrepôt EMC2
- Exigences considérées
  - Fournisseur certifié « Hébergeur de données de santé (HDS) » et conforme au RGPD
  - Offre « Infrastructure as a Service » fournissant des ressources de calcul et de stockage élastiques
  - Offre « Platform as a Service » fournissant des services logiciels entièrement gérés par l'hébergeur et intégrables dans EMC2
  - Hébergement dans des centres en France, et dans des zones de disponibilités indépendantes

I. A. Contexte du projet EMC2 (2/2)

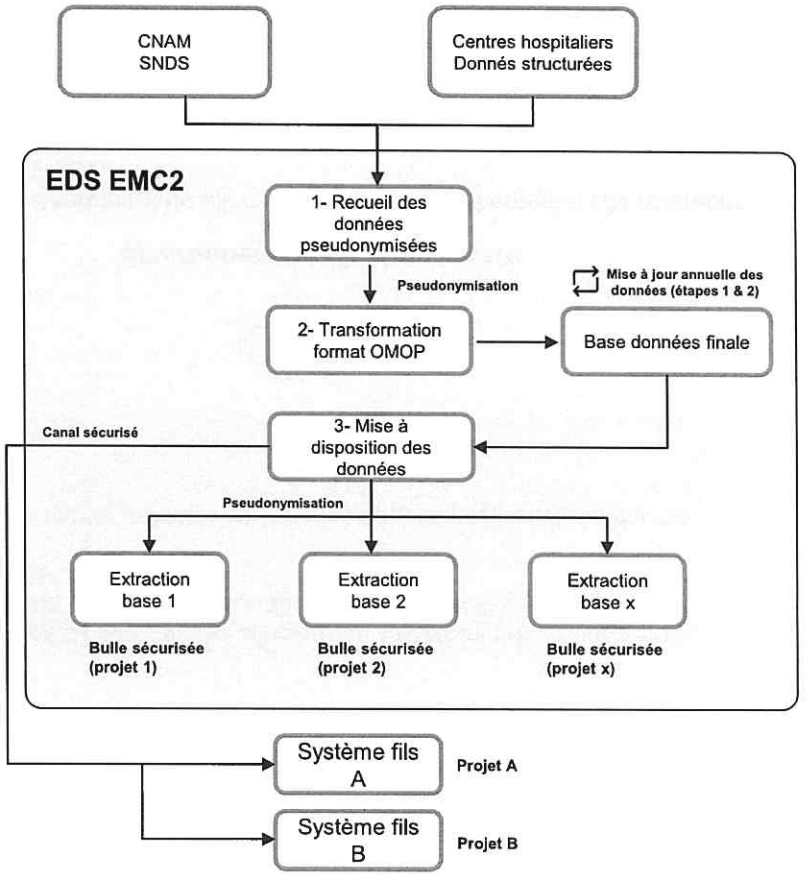
Spécificités de la plateforme EMC2

- La plateforme EMC2 permet de fournir un emplacement centralisé pour **stocker des données de santé**, selon les spécificités d'un EDS et l'exercice des droits d'un EDS :
  - Isolation forte entre les bulles sécurisées
  - Gestion des mises à jour de données (contrairement aux solutions fournissant uniquement des espaces de travail sécurisés pour l'analyse des données)
  - Système de mise à disposition de sous-ensembles de données dans des bulles sécurisées
  - Pseudonymisation à chaque transfert de données entre bulles
  - Besoin de pseudonymes déterministes pour gérer les mises à jour des données après chaque nouvel import de mise à jour des EDS ou du SNDS

- Le fonctionnement de la plateforme EMC2 suit le processus simplifié ci-dessous :
  1. **Recueil de données pseudonymisées** issues d'extractions des données des 4 centres participants et de la base principale SNDS.
  2. **Transformation des données** au format standard adopté par l'EMA (OMOP). Une opération de **mise à jour annuelle** des données doit être réalisée par le HDH
  3. **Mise à disposition des données** pour les utilisateurs selon **2 modes d'accès**
     - **Espaces de travail projet sécurisés** dans un environnement de la plateforme technologique du HDH garantissant l'intégrité de la donnée, accessible uniquement aux utilisateurs habilités (authentification forte)
     - **Transfert par canal sécurisé** en sortie de la plateforme HDH vers un système fils conforme aux prérequis de sécurité SNDS (homologation SNDS)
  4. **Décommissionnement des espaces de travail** à la fin des projets et suppression des données après une durée d'archivage de 12 mois

Les solutions fournissant des **espaces de travail sécurisés** ne gèrent **aucune opération de préparation et de mise à jour de données**. Elles fournissent uniquement des services clé en main de bulles sécurisées permettant une souplesse fonctionnelle dans un cadre stricte de sécurité, notamment de confidentialité et de traçabilité des opérations utilisateur.

Schéma EDS EMC2 simplifié



I. B. Examen par la CNIL de la demande d'autorisation du GIP PDS « Plateforme des données de santé » concernant la constitution d'un entrepôt de données de santé, dénommé "EMC2"

- En 2023, le groupement GIP PDS présente **un dossier d'autorisation à la CNIL** concernant la constitution de l'entrepôt de santé EMC2 dénommé « EMC2 » présenté en séance plénière le 12 octobre 2023 impliquant la possibilité pour le GIP PDS de disposer d'une copie partielle de la base principale du SNDS

- **Le projet de délibération CNIL** recommande le recours à **un sous-traitant qualifié SecNumCloud et donc immunisé contre tout accès non autorisé par des autorités publiques d'Etat tiers** *(Observation n° 57 – voir annexe 1)*.

- **Le Rapport** relatif à cette demande présenté en séance plénière le 12 octobre 2023 suggère de **conditionner le versement des données** au changement de l'hébergeur actuel de la PDS (Microsoft Azure) pour un hébergeur souverain, **selon 2 options** au choix du GIP PDS :
  - Option 3. Autorisation des opérations préparatoires jusqu'en 2025 et versement des données conditionné au changement d'hébergeur,
  - Option 4. Autorisation sous condition de changement d'hébergeur vers un hébergeur souverain ab initio.
- Les autres options sont l'autorisation sans condition concernant l'hébergeur (option 1), autorisation jusqu'en 2025 et changement d'hébergeur à cette date (option 2) et le refus d'autorisation (option 5).
- Le rapport précise qu'un hébergeur est souverain quand les données hébergées sont immunisées contre tout accès non autorisé par des autorités publiques d'Etat tiers.
- Le rapport recommande le recours à un hébergeur qualifié SecNumCloud et cite aussi comme solutions souveraines possibles le CASD, l'INRIA et DOCAPOSTE.

- **L'examen en séance plénière** de la demande d'autorisation CNIL du projet EMC2 conclu à :
  - Une **prorogation de l'avis** de la CNIL au 21 décembre 2023
  - Une **demande d'étude** auprès de la DNS, de la DINUM et de l'ANS de la faisabilité d'héberger l'Entrepôt de Données de Santé du projet EMC2 sur une plate-forme souveraine, résultats attendus pour le 13 décembre 2023 (option 3 ou option 4).

 * Liste des offreurs sélectionnés (voir diapositive 14)

I. C. Décision de lancement de la mission EMC2 et objectifs de la mission EMC2

- **21/10/2023 :** décision par le gouvernement de lancer la mission EMC2 à mener par la DNS, l'ANS et la DINUM
- **23/10/2023 :** lancement par la DNS, l'ANS et la DINUM de la mission EMC2
    - Pour émettre une ou plusieurs recommandations sur les sur les options 3 et 4 suggérées par la CNIL dans son rapport, la mission répondra aux questions suivantes :
        - Quelles sont les **exigences minimales** applicables au projet EMC2 en termes de réglementation, de fonctionnalités et de sécurité ?
        - Quel est le niveau de **couverture** des exigences minimales du projet EMC2 par les "hébergeurs souverains" cités dans le rapport de la CNIL ?
        - Quels sont les **impacts** pour le HDH et le projet EMC2 d'un point de vue technique, organisationnel et financier de l'hébergement de l'EDS EMC2 sur une nouvelle plateforme d'hébergement ?

 * Liste des offreurs sélectionnés (voir diapositive 14)

II. Méthodologie adoptée

a) Description de la structuration de la mission
b) Constitution de la liste des offreurs
c) Approche de construction du catalogue des exigences
d) Critères de comparaison des scénarios d'hébergement

II. A. Description de la structuration de la mission d'expertise technique (1/2)

1. Présentation des chantiers

La mission d'expertise lancée par la DNS, l'ANS et la DINUM est structurée en 4 chantiers :

- **Chantier 1 : Construction du catalogue des exigences minimales EMC2**
  - **Objectifs** : définir la liste des exigences minimales à partir des référentiels réglementaires / conformité applicables, du démonstrateur Cloud DINUM, architecture et cas d'usages EMC2
  - **Livrables** : catalogue des exigences minimales

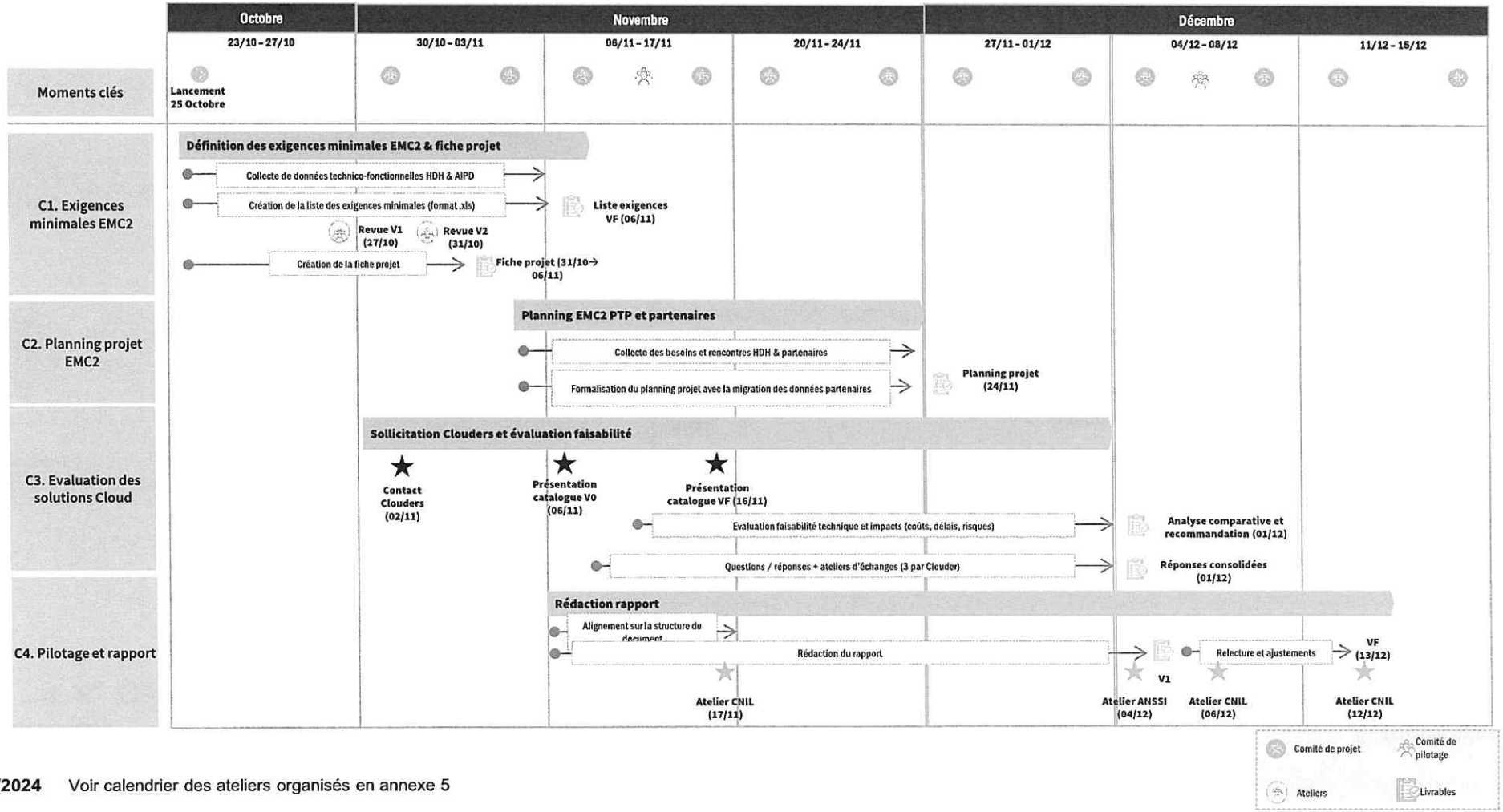
- **Chantier 2 : évaluer les impacts** (projet EMC2, techniques, organisationnels et financiers) de l'hébergement de l'EDS EMC2 sur une nouvelle plateforme d'hébergement
  - **Objectifs** : Etablir le planning projet d'EMC2 et évaluer les conséquences pour le HDH de la mise en œuvre d'une PTF supplémentaire
  - **Livrables** : Planning projet EMC2 avec les jalons clés

- **Chantier 3 : Evaluation des solutions d'hébergement**
  - **Objectifs** : Identifier les options d'hébergement EMC2 alternatives et évaluer leurs impacts
  - **Livrables** : Analyse comparative des scénarios et recommandations

- **Chantier 4 : Pilotage de la mission et production du rapport d'expertise**
  - Objectifs : Piloter l'avancement des chantiers et rédiger le rapport d'expertise à destination de la CNIL
  - Livrables : Rapports d'avancement de la mission; Rapport d'expertise technique

II. A. Description de la structuration de la mission d'expertise technique (2/2)

2. Calendrier de la mission par chantier



 Voir calendrier des ateliers organisés en annexe 5

II. A. Conditions de réalisation de l'étude

Une mobilisation très active des différentes parties prenantes tout au long de la mission

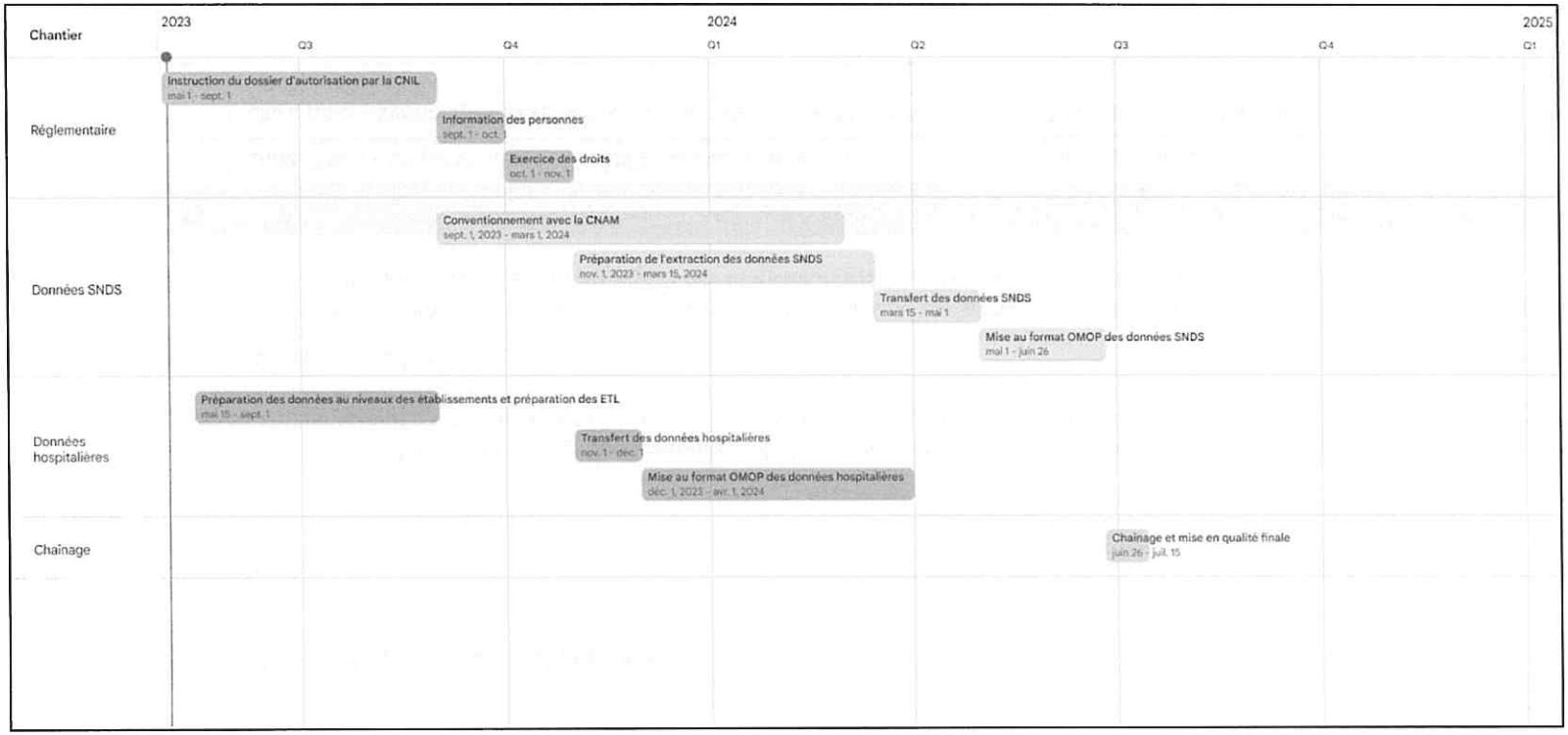
- Une forte réactivité et une grande agilité des offreurs
- Une forte implication et collaboration du HDH
- L'expertise de l'ANSSI, de l'INRIA et de la CNIL sur les exigences applicables

Acteurs	Participation à la mission EMC2
HDH	• Participation aux échanges avec les offreurs • Apport d'expertise technico-fonctionnelle nécessaire à la définition d'exigences adaptées aux besoins
INRIA	• Apport d'expertise et participation avec l'équipe de la mission EMC2 à l'évaluation des réponses des fournisseurs aux exigences
Cloud Temple **Outscale/Numspot/HEVA** **OVH/Cleyrop** **Clinityx** **CASD**	• Participation aux séances plénières • Participation aux entretiens bilatéraux • Formalisation des réponses au catalogue d'exigence EMC2
ANSSI	• Apport d'expertise en lien avec les exigences de sécurité et l'évaluation des feuilles de route fournisseurs
CNIL	• Apport d'expertise en lien avec les exigences réglementaires du référentiel EDS et RGPD

II. A. Calendrier prévisionnel de mise en œuvre de l'EDS EMC2

Les engagements du projet EMC2 vis-à-vis de l'EMA portent sur les jalons suivants, pour une mise en production à mi-Juillet 2024 :

- Validation de la conformité réglementaire de l'EDS - CNIL
- Préparation et extraction des données SNDS / établissements de santé
- Mise au format OMOP des données
- Chainage et mise en qualité finale



II. B. Constitution de la liste des offreurs

La liste des offreurs interrogés est constituée de :

1 Hébergeurs qualifiés SecNumCloud

Fournisseurs disposant de la qualification ANSSI SecNumCloud, comprenant également l'immunité contre toute réglementation extraterritoriale. Ce critère, en ligne avec la doctrine « Cloud au centre », permet de ne retenir que des fournisseurs répondant à une définition explicite de la notion de souveraineté (critère 19.6 de SecNumCloud)

2 Entrepôts de données de santé validés par la CNIL

Les fournisseurs d'entrepôts de données de santé recommandés par la CNIL suite aux conclusions du projet de délibération en réponse à la demande d'autorisation du projet EMC2. Après analyse des EDS avec appariement de données de santé au SNDS par la mission EMC2, Clinityx a été ajoutée à la liste des offreurs.

Liste des offreurs	Description	Participation à l'étude
Oodrive	Offre SaaS - Oodrive_meet – Qualifié SecNumCloud 3.2	Non, décision Oodrive
Worldline	Offre IaaS - Worldline Cloud Services - Secured IaaS- Qualifié SecNumCloud 3.1	Non, décision Worldline
OVH	Offre IaaS – Private Cloud - Qualifié SecNumCloud 3.2	Oui
Cloud Temple	Offre Iaas – Secure Temple – Qualifié SecNumCloud 3.1	Oui
Numspot / Outscale / Heva	Offre IaaS - Clloud on Demand – Outscale - Qualifié SecNumCloud 3.2	Oui
Clinityx	Autorisation de traitement CNIL de projet EDS avec appariement	Oui
CASD	Autorisation de traitement CNIL de projet EDS avec appariement	Oui
Inria	Espace de travail pour les chercheurs Arcana hébergé par outscale	Non, l'INRIA a alors été associée à l'équipe menant la mission EMC2. En effet, l'INRIA a considéré que leur offre n'avait pas vocation à répondre au besoin du projet EMC2

II. C. Approche de construction du catalogue des exigences (1/2)

La construction du catalogue des exigences minimales a été réalisée en 3 étapes et en repartant de 3 sources d'exigences :

- Référentiels EDS & SNDS, PGSSI-S, PSSI MCAS, RGS
- Documents d'architecture EMC2
- Démonstrateur Cloud DINUM

1- Construction des exigences réglementaires

- Identification des exigences unitaires incluses dans les référentiels EDS et SNDS
- Rapprochement et fusion, s'il y a lieu, des exigences de conformité, sécurité et réglementaires issues des deux référentiels
- Sélection des exigences minimales applicables à un hébergeur Cloud
- Intégration des exigences de respect HDS et SNC

2- Construction des exigences techniques spécifiques EMC2

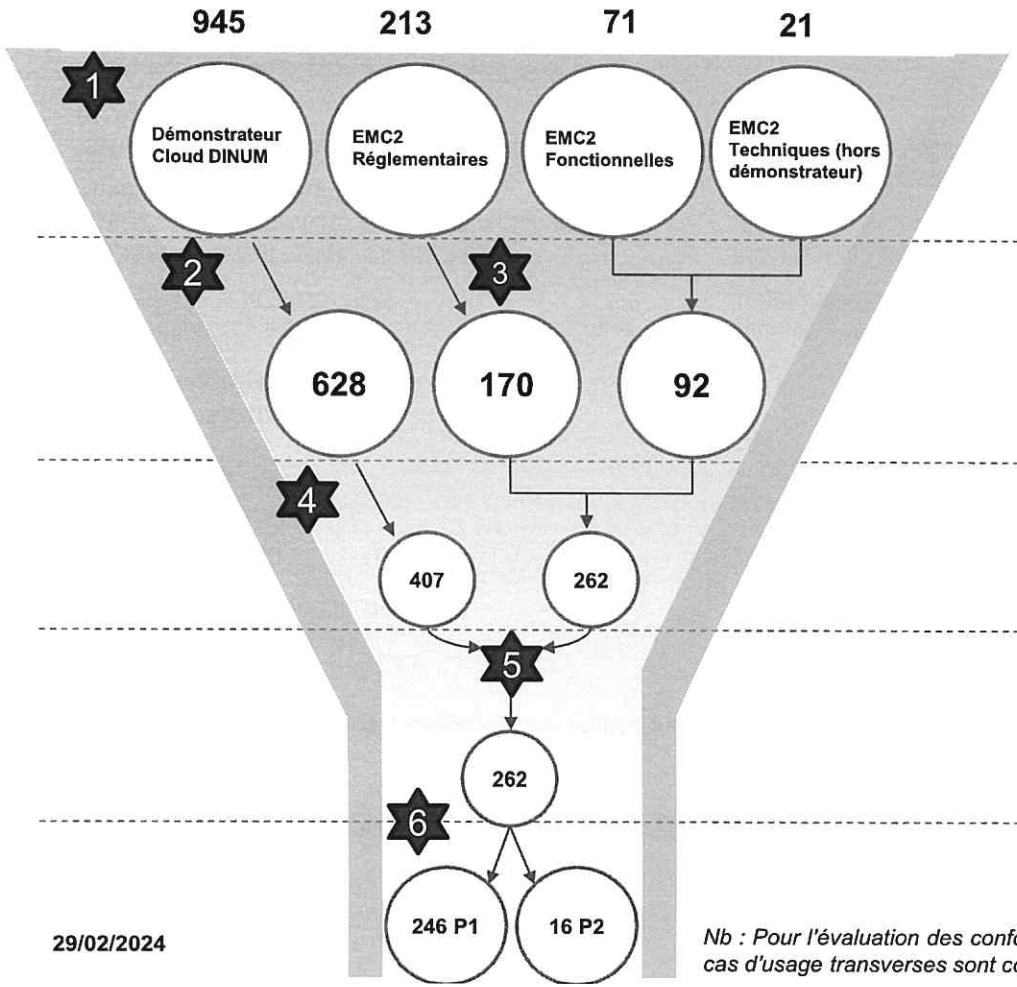
- Définition de la liste des exigences fonctionnelles de la plateforme EMC2
- Identification des briques de services techniques répondant aux exigences fonctionnelles
- Etablissement de la liste des exigences fonctionnelles et techniques relatives aux spécificités de la plateforme EMC2

3- Rapprochement avec les exigences du démonstrateur Cloud DINUM

- Rapprochement de la liste des activités du démonstrateur DINUM (statut « Must » et « Should » de HDH) avec la liste des exigences réglementaires, techniques et fonctionnelles déjà identifiées
- Pour les activités non rapprochées, évaluation de la pertinence de leur applicabilité comme exigences minimale pour EMC2, ajout le cas échéant à la liste des exigences techniques déjà établie

La conformité avec les réglementations de la CNIL, y compris les éléments permettant de produire une AIPD, est une exigence légale. Elle garantit qu'EMC2 respecte les règles du RGPD et protège efficacement les données
Les retours de la CNIL du 04/12 ont été analysés pour mettre à jour le catalogue EMC2 (en annexe 4)

II. C. Approche de construction du catalogue des exigences (2/2)



Étape 1 : Identification des exigences sources à considérer pour EMC2

- Démonstrateur Cloud DINUM (945 exigences techniques)
- Référentiels EDS, SNDS, PGSSI-S, PSSI MCAS, CNIL, ANSSI (180 exigences réglementaires)
- Architecture technico-fonctionnelle EMC2 (73 exigences fonctionnelles et 19 exigences techniques hors démonstrateur

Étape 2 : Sélection des exigences du démonstrateur Cloud fourni par HDH (M et S) **(-317)**

Étape 3 : Suppression des doublons des différents référentiels réglementaires. **(-41)**

Étape 4 :

- Réduction des exigences minimales, en concertation avec HDH **(- 221 dont 149 BDaaS)**
- Priorisation des exigences, pour garder seulement les exigences minimales au fonctionnement du projet EMC2

Étape 5 : Mapping et regroupement des exigences techniques du démonstrateur avec les exigences réglementaires et fonctionnelles

Étape 6 : Priorisation des exigences entre P1 (Must) et P2 (Nice to have)

Nb : Pour l'évaluation des conformités, les exigences informatives sont exclues de la notation et les cas d'usage transverses sont comptabilisés une seule fois (annulation de l'effet doublon)

Numéro de l'étape

II. D. Critères de comparaison des scénarios d'hébergement (1/2)

Deux scénarios identifiés comme potentielles options d'hébergement alternatives pour le projet EMC2 :

Scénario	Description	Offreurs considérés
Scénario 1 : Construction d'une nouvelle plateforme d'infrastructure basée sur des services qualifiés SecNumCloud avec réutilisation des briques développées par HDH	**Caractéristiques clés** : • Migration vers une nouvelle infrastructure IaaS qualifiée SecNumCloud. • Intégration des briques logicielles existantes de HDH • Usage de briques logicielles déjà existantes et utiles pour HDH • Possibilité d'utiliser des services managés : Kubernetes, KMS, HSM, etc.	Cloud Temple, OVH/Cleyrop, Outscale/Numspot/HEVA
Scénario 2 : Mise à disposition d'une plateforme incluant tous les services applicatifs nécessaires au projet EMC2	**Caractéristiques clés** : • Plateformes déjà validées par la CNIL • Intégration des spécificités du HDH • Solution définitive dans le cadre d'une plateforme IaaS possédant une feuille de route de qualification SNC ou construite par une entité publique	Clinityx, CASD

II. D. Critères de comparaison des scénarios d'hébergement (2/2)

Les deux scénarios considérés seront comparés sur la base des familles de critères suivantes :

Les exigences minimales du catalogue de conformité sur les volets technique, réglementaire et fonctionnel

Les spécificités et prérequis relatifs au projet EMC2 et aux engagements pris vis-à-vis de l'EMA : le respect de la date de mise en production définie au 15 Juillet 2024, l'existence du véhicule d'achat des services, l'impact des coûts d'adoption des services (migration, mise en conditions opérationnelles/sécurité, usage des services)

L'indice de confiance d'évolution des services et sa cohérence avec la feuille de route du projet EMC2 et du HDH : Evalué sur la base des preuves documentaires fournies par les offreurs, les feuilles de routes déclaratives de disponibilité des services partagées par les offreurs dans le cadre du démonstrateur Cloud au centre et la présente étude, les délais de qualification SecNumCloud (avis de l'ANSSI)

Avertissement

Les **délais courts** de la mission n'ont pas permis de :

- Vérifier les réponses des offreurs aux exigences par des tests, l'évaluation a été réalisée sur la base des preuves fournies, des documents publics disponibles et d'un indice de confiance apprécié par l'équipe projet
- Réaliser le chantier 2, et donc une analyse fine des impacts des différents scénarios sur le HDH
- Effectuer une évaluation financière sur une base de coûts comparables, les réponses des offreurs étant formalisées avec des logiques et mécanismes différents

Les taux de couverture ne sont que pour les **besoins priorisés du projet EMC2** et ne doivent pas guider le choix d'un fournisseur pour un autre projet.

III.Évaluation des scénarios d'hébergement

a) Description de l'approche d'évaluation
b) Analyse comparative des scénarios
c) Bilan de comparaison des offres

III. A. Description de l'approche d'évaluation

Scénario 1 : Approche d'évaluation

1 Analyse des réponses fournisseurs

- ✓ **Pour les exigences déclarées couvertes par les fournisseurs évalués** : Évaluation des déclarations de conformité des fournisseurs et vérification de cette conformité au travers de preuves (documentations publiques, références projets, expérience sur le sujet, etc.)
- ⓘ **Pour les exigences déclarées prochainement couvertes par les fournisseurs évalués (roadmap déclarative entre 3 et 6 mois)** : Elaboration d'un indice de confiance basé sur l'analyse des preuves fournies, mises en perspective avec les connaissances du marché
- ✕ **Pour les exigences indiquées comme problématiques par les fournisseurs** : Analyse des solutions de contournement proposées et de leurs impacts par catégorie de services
- Sur la base des analyses, réalisation d'une **notation binaire** (0 : non conforme, 1 : conforme) et calcul d'un pourcentage de conformité par offreur à 3 temps (à date, 3 mois, 6 mois).

2 Analyse comparative des scénarios

- **Pour chaque scénario ;**
  - Présentation des réponses offreurs (Scénario 1) ou présentation des offres (Scénario 2)
  - Evaluation des non-conformités sur 3 axes :
    - Axe sécurité : impact sécurité des non-conformités techniques, solutions de contournement (impacts et faisabilité), absence qualification SecNumCloud pour certains services
    - Axe fonctionnel : Impacts fonctionnels des non-conformités techniques, solutions de contournement (impacts et faisabilité)
    - Axe projet : Délais auto-homologation SNDS / Qualification SecNumCloud / Véhicule d'achat
  - Bilan de chaque scénario
  - Limitation : Pas d'évaluation des coûts
- **Synthèse comparative des scénarios**

III. B. Analyse comparative des scénarios

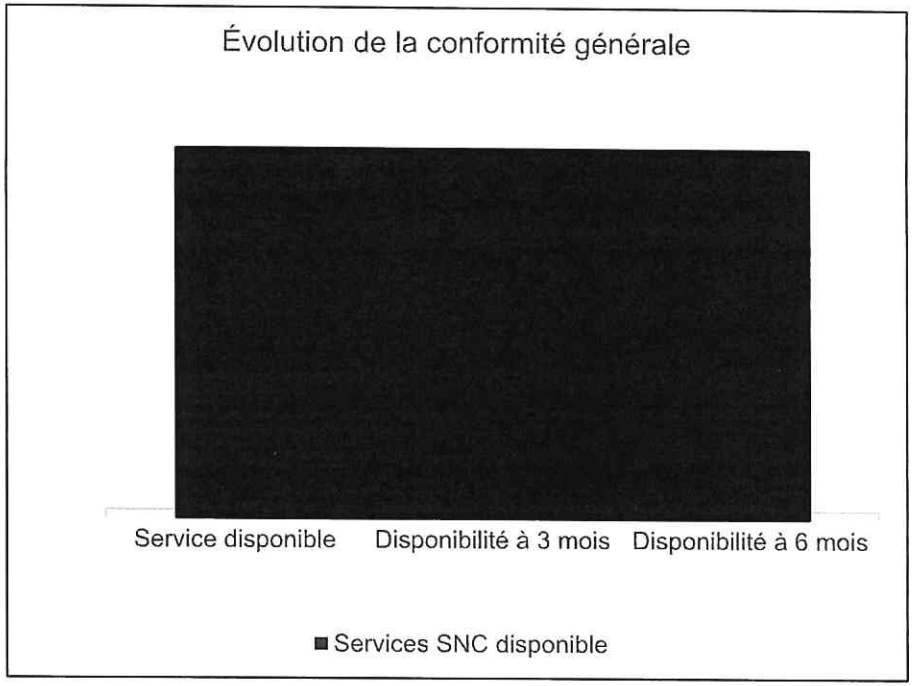
Scénario 1 : Conformité au catalogue d'exigences minimales EMC2 - OVH / Cleyrop

Description de l'offre

- OVH Cloud est le leader Français du secteur en termes de chiffre d'affaires et dispose de trois offres de Cloud : Public Cloud, Bare Metal et VMware on OVHCloud
- L'offre qualifiée SecNumCloud repose sur VMware on OVHCloud et utilise une solution propriétaire commercialisée par Broadcom
- Cleyrop est une entreprise française proposant une solution unifiée pour gérer et industrialiser l'usage des données.
- [redacted]
- [redacted]

Écarts entre taux déclaratif et estimé

[redacted]



Conformité	Taux d'exigences conformes déclaratif	Taux d'exigences conformes estimé	Briques conformes SNC
A date	[redacted]		
Roadmap 3 mois			
Roadmap 6 mois			

Type d'exigence	Fonctionnelle	Réglementaire	Technique
A date	[redacted]		
Roadmap 3 mois			
Roadmap 6 mois			

III. B. Analyse comparative des scénarios

Scénario 1 : Conformité au catalogue d'exigences minimales EMC2 – Outscale / Numspot / Heva

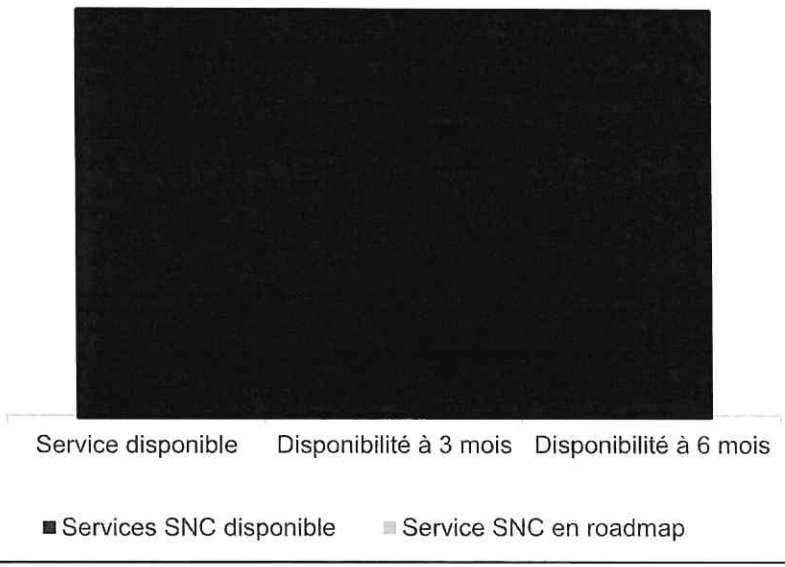
Description de l'offre

- Outscale est une filiale de Dassault Systèmes qui fournit un IaaS déjà qualifié SNC
- Numspot est une société commune entre Docaposte, la Banque de Territoires, Dassault Systèmes et Bouygues Telecom. Elle développe des services PaaS reposant sur le socle IaaS existant d'Outscale, en vue d'une qualification SNC
- Le groupement utilise un mix de solutions internes, commerciales et de solutions open source
- L'offre étudiée Outscale/Numspot/HEVA prévoit une évolution significative des offres de services SNC pour 2024
- [redacted]

Écarts entre taux déclaratif et estimé

- [redacted]

Évolution de la conformité générale



Conformité	Taux d'exigences conformes déclaratif	Taux d'exigences conformes estimé	Briques conformes SNC
A date	[redacted]		
Roadmap 3 mois			
Roadmap 6 mois			

Type d'exigence	Fonctionnelle	Réglementaire	Technique
A date	[redacted]		
Roadmap 3 mois			
Roadmap 6 mois			

III. B. Analyse comparative des scénarios

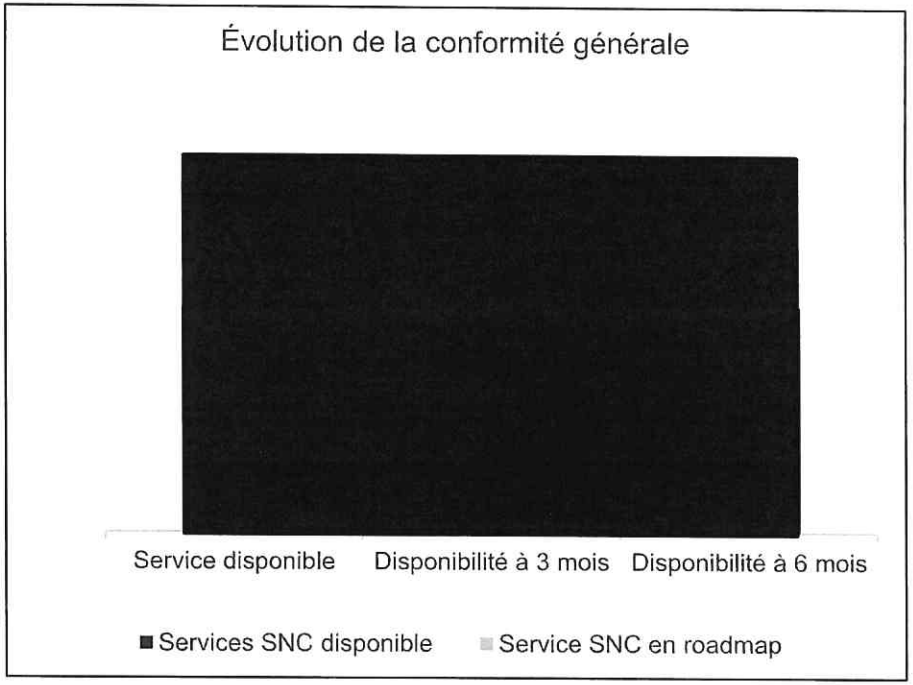
Scénario 1 : Conformité au catalogue d'exigences minimales EMC2 – Cloud Temple

Description de l'offre

- Cloud Temple est une filiale du groupe Neurones dédiée à l'infogérance et à l'hébergement.
- Cloud Temple fournit un IAAS qualifié SNC et souhaite étendre le catalogue de services à des services plus évolués.
- [redacted]

Écarts entre taux déclaratif et estimé

- [redacted]



Conformité	Taux d'exigences conformes déclaratif	Taux d'exigences conformes estimé	Briques conformes SNC
A date	[redacted]		
Roadmap 3 mois			
Roadmap 6 mois			

Type d'exigence	Fonctionnelle	Réglementaire	Technique
A date	[redacted]		
Roadmap 3 mois			
Roadmap 6 mois			

III. B. Analyse comparative des scénarios

Scénario 1 : Bilan de la conformité des offreurs

L'évaluation des réponses aux exigences EMC2 révèle une bonne couverture des fournisseurs Cloud sur les briques d'hébergement de base, avec une ambition d'étendre leur empreinte de qualification à des services avancés :

	SecNumCloud	Réseau	Calcul (élasticité)	Stockage (bloc & objet)	IAM (Granularité fine)	Kubernetes	Traces	SIEM	Chiffrement (KMS+ HSM)
OUTSCALE / NumSpot / OVHcloud / CLOUD TEMPLE	[redacted]								
Si non conforme, solution de contournement proposée		Utilisation de Firewall/VLAN pour la segmentation réseau	Dimensionnement sur les pics d'activité	Utilisation d'une solution tierce (pour le chiffrement)	Création de tenant par projet	Installation et gestion de Kubernetes par le HDH	Réalisation de contrôles manuels	Utilisation d'une solution tierce	Utilisation d'une solution tierce

Légende: **Conforme** (service SNC disponible) — **Partiellement conforme** (une partie des services SNC sont disponibles, le reste est en prévision/roadmap 6 mois) — **Non conforme** (non disponible sous 6 mois)

- ✓ Haute disponibilité et redondance des infrastructures d'hébergement
- ✓ Capacités de gestion centralisée de l'écosystème d'hébergement
- ✓ Capacités de sécurité et du contrôle d'accès aux infrastructures et à la donnée

III. B. Analyse comparative des scénarios

Scénario 1 : Axe sécurité (1/2)

	Si pas disponible	Solution de contournement	Impacts de la solution de contournement	Complexité opérationnelle
Bastion	• **Risque d'attaques directes** en l'absence d'un point de contrôle d'accès centralisé, permettant de contenir la compromission des données sensibles au niveau de l'ensemble des bulles • **Difficulté de surveillance** de l'accès distant et des activités des utilisateurs	• Intégration d'un bastion dans l'architecture de la solution EMC2, devant être administrée par le HDH	• Coûts et délais relatifs à la mise en place et la gestion d'un bastion par le HDH	Faible
Chiffrement (HSM / KMS)	• **Exposition des clés cryptographiques** : sans un HSM, les clés cryptographiques peuvent être plus vulnérables aux attaques, compromettant l'intégrité et la confidentialité des données. • **Difficulté de conformité au RGPD** : un HSM fournit un environnement sécurisé spécifique à la gestion des clés, renforçant la protection contre les accès non autorisés. • **Complexité de gestion** de cycle de vie des clés de chiffrement (+100 clés nécessaires pour assurer la sécurité de la donnée)	• Gestion manuelle des clés de chiffrement au travers d'un stockage sécurisé *3 types de clés à gérer dans le contexte EMC2 : chiffrement des données (repos, objet), chiffrement pour import de clés depuis une solution externe, clés de pseudonymisation*	• Double complexité opérationnelle : gestion du cycle de vue des clés via la mise en place d'une politique de chiffrement; mise en œuvre d'une solution de stockage sécurisé des clés accompagnée d'une politique d'accès	Très complexe
IAM (avec forte granularité)	• **Droit d'accès excessif aux ressources et données** : L'absence de granularité dans la gestion des identités peut conduire à des autorisations excessives, où les utilisateurs peuvent avoir accès à des données sensibles au-delà de leur rôle. • **Difficulté de gestion des accès** : La gestion des droits d'accès peut devenir complexe, avec un risque accru d'erreurs humaines d'attribution de droits pouvant conduire à des failles de sécurité.	• Création de tenants par projet : gestion segmentée des accès ressources au niveau du projet	• Complexité de mise en place (très dépendante de l'architecture) et de gestion (niveau d'isolation, ségrégation de la donnée, scalabilité, décommissionnement, conformité aux politiques d'usage et de sécurité) • 9 types de profils / rôles définis pour EMC2	Très Complexe

III. B. Analyse comparative des scénarios

Scénario 1 : Axe sécurité (2/2)

	Si pas disponible	Solution de contournement	Impacts de la solution de contournement	Complexité opérationnelle
Réseau (Micro-segmentation)	• **Risque d'accès au stockage** depuis l'ensemble des bulles sécurisées au sein de l'entrepôt de données • **Détection plus complexe et moins rapide** des comportements anormaux (supervision plus large)	• Mise en place de firewall pour chaque service d'infrastructure (stockage, VDI, bastion, serveurs de logs, conteneurs, etc.) → **100 firewalls environ** pour EMC2	• Complexité de gestion des configurations (scalabilité, conformité, tests, ..), délai de mise en œuvre, risque d'erreurs des configurations	Très Complexe
SIEM	• **Détection tardive des activités suspectes** ou des violations de sécurité, compromettant la réactivité face aux menaces • **Manque de visibilité sur les événements critiques** rendant difficile la compréhension complète des activités de sécurité et des relations de cause à effet entre différents événements • **Difficulté dans l'investigation** et réduction de la capacité à enquêter sur des incidents de sécurité et à comprendre leur impact (perte de la puissance des algorithmes d'analyse et de centralisation de la donnée) • **Difficulté dans la gestion et la configuration** des politiques de sécurité	• Mise en place de mécanismes de journalisation étendue au niveau des systèmes individuels pour collecter des journaux d'événements. • Gestion et analyse manuelle des traces et exposition graphique pour analyse	• Délai d'intégration et de configuration • Besoin de ressources d'infrastructures adaptées pour la gestion de larges volumes de données de journalisation (semi-structurées)	Complexe
Traces (vérification de conformité)	• Complexité de valider la conformité de la plateforme aux politiques définies d'une manière automatisée et précise	• Réalisation de contrôles manuels	• Risques de non-conformité (notamment sur les aspects sécurité) • Délais conséquents de traitement et mobilisation de plusieurs ressources humaines	Complexe

III. B. Analyse comparative des scénarios

Scénario 1 : Axe fonctionnel

	Si pas disponible	Solution de contournement	Impacts de la solution de contournement	Complexité opérationnelle
Stockage Objet	• Modification de l'architecture logicielle EMC2	• Utilisation d'autres services de stockage (ex. HDFS) • Utilisation d'une solution de stockage logiciel (ex : Minio)	• Changement dans le modèle de stockage et la gestion des données stockées • Management d'une solution tierce	Faible
Kubernetes	• Gestion d'un cluster Kubernetes par le HDH et/ou revue de l'architecture logicielle des solutions du HDH.	• Installation, configuration et gestion d'un cluster Kubernetes par le HDH	• Complexité d'installation et de mise en condition opérationnelle (gestion de la disponibilité, mises à jour, évolutivité des ressources) • Montée en compétences des équipes du HDH en charge et en couverture horaire de la MCO.	Complexe
Calcul (Non-scalabilité des capacités de calculs)	• Risque de retards de traitement des projets lors des pics de charge bi-annuels (mise à jour SNDS et mise au format OMOP) • Délais importants de réexécution de traitements infructueux	• Surdimensionnement des infrastructures afin d'absorber les pics de charge bi-annuels	• Usage non optimisé des capacités infrastructures disponibles • Augmentation des coûts • Perte d'agilité et de flexibilité dans les réponses aux besoins	Complexe

III. B. Analyse comparative des scénarios

Scénario 1 : Axe projet

1. Liaisons SNDS / Centres hospitaliers

- **Opérations techniques de raccordement** entre la nouvelle plateforme et le SNDS / Centres hospitaliers
  - Reconfiguration des canaux d'accès (VPN – adaptation des adresses IP)
  - Repartage des clés de chiffrement des communications pour garantir la sécurité des données échangées
- **Mise à jour des conventions** avec la CNAM et les centres hospitaliers
- Impact calendrier EMC2 **environ 1 à 3 mois**

2. Homologation SNDS

- **Opérations d'auto-homologation** à réaliser par le HDH :
  - **Instruction d'un document d'architecture technique (DAT)**
  - Analyse de risque et mesures de remédiation
  - Audit PASSI de la nouvelle plateforme
  - Vérification de la conformité aux référentiels PSSI-MCAS, EDS et RGS
  - Signature d'une convention CNAM
- Impact calendrier EMC2 de **10 à 12 mois**

3. Impact organisationnel HDH

- **Formation des équipes** HDH aux nouvelles technologies à adopter
- **Appropriation des spécificités technologiques** en termes d'architecture et de sécurité
- Adaptation du modèle opérationnel pour intégrer les opérations MCO & MCS

➢ Quel que soit le fournisseur de services IaaS SecNumCloud choisi pour l'hébergement de l'EDS EMC2, une **adaptation des liaisons SNDS/Centre Hospitaliers et homologation SNDS sera nécessaire** pour opérer en environnement cible

➢ Au global, ces opérations introduisent un **impact planning d'environ 12 mois**, hors délais de couverture des non-conformités (mise en œuvre de contournements ou mise à disposition de nouveaux services SNC), délai achat, d'autorisation CNIL (AIPD) et configuration de la solution

III. B. Analyse comparative des scénarios

Scénario 2 : Approche d'évaluation

1 Analyse des réponses fournisseurs (scénario 2)

Les acteurs retenus dans le cadre du scénario 2 ont déjà mis en œuvre des EDS validé par la CNIL, l'analyse consiste à vérifier s'ils sont en capacité de proposer des services au HDH afin de réponde aux besoins d'EMC2.

L'analyse a été conduite selon ces critères :

- **Conformité fonctionnelle :** Vérifier que la plateforme répond aux exigences fonctionnelles définies par HDH pour EMC2
- **Couverture des services disponibles :** Vérifier que la plateforme offre les services nécessaires pour EMC2.
- **Mesures de sécurité :** Vérifier que la plateforme répond aux exigences de sécurité et de ségrégation d'un entrepôt de santé
- **Flexibilité et adaptation de l'offre aux besoins :** Vérifier que la plateforme peut être adaptée aux besoins spécifiques de EMC2.

2 Analyse comparative des scénarios

- **Pour chaque scénario ;**
  - Présentation des réponses offreurs (Scénario 1) ou présentation des offres (Scénario 2)
  - Evaluation des non-conformités sur 3 axes :
    - Axe sécurité : impact sécurité des non-conformités techniques, solutions de contournement (impacts et faisabilité), absence qualification SecNumCloud pour certains services
    - Axe fonctionnel : Impacts fonctionnels des non-conformités techniques, solutions de contournement (impacts et faisabilité)
    - Axe projet : Délais auto-homologation SNDS / Qualification SecNumCloud / Véhicule d'achat
  - Limitation : Pas d'évaluation des coûts
- **Synthèse comparative des scénarios**

MINISTÈRE
DE LA SANTÉ
ET DE LA PRÉVENTION
Liberté
Égalité
Fraternité

III. B. Analyse comparative des scénarios

Scénario 2 : CASD - Bilan de conformité de l'offre

Délégation ministérielle au numérique en santé

Présentation de l'offre CASD

- Le CASD offre un service clé en main de bulles sécurisées
- La SD-BOX (Terminal utilisateur autonome sécurisé fourni par le CASD) est indispensable pour accéder aux données
- HDH ne se charge que de définir ses besoins fonctionnels
- Le contrat type du CASD permet de définir les responsabilités, les modalités de service, qualité de service
- HDH peut disposer de bulles dédiées pour la gestion des identifiants, la préparation des données et transfert ensuite les données aux bulles projets
- Possibilité de définir des outils spécifiques dans les bulles (python, spark, R, duckdb, Atlas…)

Répartition des services proposés : Proposition d'architecture

[redacted]

Détails de la conformité

- **Taux d'exigences conforme déclarative à date** : [redacted]
- **Limitations** : [redacted]
- **État de la qualification SecNumCloud** : N/A (entité publique)
- **Délais avant la qualification SecNumCloud** : N/A (entité publique)
- **Indice de confiance roadmap SNC** : N/A (entité publique)

Planning de mise à disposition de l'environnement CASD (hors migration EMC2)

[redacted]

III. B. Analyse comparative des scénarios

Scénario 2 : CASD

Axe Sécurité

- En réponse aux besoins d'EMC2, le CASD propose de développer une solution 'sur-mesure' sur la base de briques existantes. [redacted]
- [redacted]
- [redacted]
- [redacted]

Axe Fonctionnel

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

Axe Projet

- Concernant le cadre contractuel, une coopération entre le HDH et le CASD doit être mise en place en respectant le modèle de collaboration public/public (Article 2511), faute de quoi il pourrait y avoir un risque de requalification en marché public et un risque pénal.*
- Le choix du CASD comme solution d'hébergement par le HDH pourrait représenter un écart par rapport à la stratégie "Cloud au Centre" de l'État français qui favorise les environnements Cloud qualifiés SecNumCloud
- Le délai de mise à disposition de la plateforme [redacted] est un facteur-clé à considérer. Le temps nécessaire pour développer, mettre en place et tester cette plateforme affecterait le calendrier de migration du HDH [redacted] influençant significativement le déroulement du projet EMC2.

(*) Voir synthèse analyse juridique en annexe 3

III. B. Analyse comparative des scénarios

Scénario 2 : Clinityx - Bilan de conformité de l'offre

Présentation de l'offre Clinityx

- Clinityx a été autorisé par la CNIL à mettre en œuvre 5 entrepôts de données, accessibles uniquement par les équipes Clinityx, alimentés en toute ou partie par des données du SNDS.
- Les entrepôts de données de Santé de Clinityx sont hébergés entièrement depuis Septembre 2023 chez Cegedim.Cloud qui est un hébergeur certifié HDS, ISO27001

Répartition des services proposés : Proposition d'architecture

Détails de la conformité

- **Taux d'exigences conforme déclarative à date:**
- **Limitations :**
- **État de la qualification SecNumCloud** : En cours
- **Délais avant la qualification SecNumCloud :**
- **Indice de confiance roadmap SNC :**

Planning de mise à disposition de l'environnement Clinityx (hors migration EMC2)

III. B. Analyse comparative des scénarios

Scénario 2 : Clinityx

Axe Sécurité

- Suite à notre consultation, [redacted]
- Les architectures EDS de Clinityx, hébergées par Cegedim.cloud, [redacted]
- [redacted]
- [redacted]
- [redacted]

Axe Fonctionnel

- [redacted]
- [redacted]
- [redacted]

Axe Projet

- [redacted]
- [redacted]
- [redacted]

III. C. Bilan de comparaison des offres

		Couverture des exigences EMC2	Solutions de contournement des non-conformités	Impacts projet EMC2	Risques d'adoption
Scénario 1	OUTSCALE / NumSpot Cloud. Européen. Citoyen.	**Conformité:** [redacted] **Briques non-conformes:** [redacted] **Indice de confiance roadmap SNC :** [redacted]	[redacted]	• Auto-homologation : 10 à 12 mois minimum • Raccordement SNDS / établissement : 1 à 3 mois • [redacted] • Pénalités financières en cas de non respect du planning communiqué à l'EMA • Déclaration de modification de sous-traitant à l'EMA • Sous-traitance à déclarer dans la demande d'autorisation EDS, aucune action vis-à-vis de l'EMA	• Capacité du HDH à monter en compétences et maitriser les technologiques et services proposés • Capacité du HDH à gérer deux plateformes technologiques en parallèle pour une durée prolongée (dépendance délais de qualification SNC)
	OVHcloud	**Conformité:** [redacted] **Briques non-conformes:** [redacted] **Indice de confiance roadmap SNC** [redacted]			
	CLOUD TEMPLE	**Conformité:** [redacted] **Briques non-conformes:** [redacted] **Indice de confiance roadmap SNC** [redacted]	[redacted]		
Scénario 2	CISD	**Conformité fonctionnelle:** [redacted] **Limitations :** [redacted] **Indice de confiance roadmap SNC** : N/A (entité publique)	• [redacted] • [redacted]	• Auto-homologation SNDS existante, auto-homologation <1 mois • Raccordement SNDS / établissements : 1 à 3 mois • Délais de mise en œuvre plateforme : 3 mois (hors limitations et couverture de risque) • Pénalités financières en cas de non respect du planning communiqué à l'EMA • Sous-traitance à déclarer dans la demande d'autorisation EDS, aucune action vis-à-vis de l'EMA	• [redacted] • Coopération public/public □ Risque de requalification en marché public
	Clinityx / cegedim.cloud	**Conformité fonctionnelle:** [redacted] **Limitations :** [redacted] **Indice de confiance roadmap SNC :** [redacted]	• [redacted] • [redacted]	• Auto-homologation SNDS existante, auto-homologation >1 mois • Raccordement SNDS / établissements : 1 à 3 mois • Délais de mise en œuvre plateforme : 3 mois (hors limitations et couverture de risque) • [redacted] • Pénalités financières en cas de non respect du planning communiqué à l'EMA • Sous-traitance à déclarer dans la demande d'autorisation EDS, aucune action vis-à-vis de l'EMA	• [redacted]

IV. Synthèse & Recommandations

IV. Synthèse

1. La relative **indépendance** de la plateforme du HDH vis-à-vis des offres spécifiques Azure comme les bases de données managées (relationnelle, clé/valeur, document), le VDI, le DNS ou l'usine logicielle, **simplifie la portabilité**
2. La présente étude a permis une forte mobilisation de **l'écosystème SecNumCloud** avec des feuilles de route ambitieuses pour répondre aux besoins du projet EMC2 et plus globalement du HDH.
3. Les délais supplémentaires (supérieurs à 12 mois) pour mettre en œuvre une nouvelle plateforme spécifique pour le projet EMC2 entraineraient des **pénalités financières** et **un risque de non renouvellement** pour la suite du projet pour le HDH mais également une **dégradation de son image vis-à-vis de l'EMA** ce qui pourrait compromettre les futures collaborations notamment Darwin et EHDS. Enfin, il semble utile de rappeler que l'EMA a l'habitude de travailler avec de nombreux partenaires européens dont la majorité sont capables de fournir les données sous quelques mois. La candidature du HDH à l'appel d'offre de l'EMA entendait remettre la France dans la course du partage des données de santé au niveau européen
4. Les partenaires du projet EMC2 **ont déjà engagés des moyens** pour préparer ce projet, le retard et/ou le changement de trajectoire aurait un **impact négatif sur eux**, mais également sur l'image du HDH vis-à-vis de ces partenaires
5. L'adoption d'une nouvelle plateforme nécessite une **montée en compétences des équipes du HDH**, ce qui implique des risques en termes de délai et un impact sur le maintien en condition opérationnelle
6. La mission du HDH est de garantir l'accès aisé et unifié, transparent et sécurisé, aux données de santé pour améliorer la qualité des soins et l'accompagnement des patients **via une plateforme unique**. Construire une plateforme ad hoc pour un projet spécifique (comme EMC2) **peut complexifier la réalisation de cette mission**
7. La gestion de données sensibles dans une **plateforme mutualisée** implique des **exigences techniques supérieures** par exemple une granularité fine des droits IAM, la gestion automatisée des clés de chiffrement, la ségrégation réseau fine etc…
8. La **mutualisation des projets et des activités** (mise au format OMOP, ingestion de grandes volumétries de données) sur une plateforme nécessite **une forte élasticité** sur les ressources pour gérer les pics de charge
9. La doctrine cloud au centre de l'Etat **impose depuis le 5 Juillet 2021**, sauf dérogation par la Première Ministre, **pour tout nouveau projet** de système d'information d'utiliser des services **qualifiés SecNumCloud** pour les applications manipulant des données sensibles

IV. Recommandations

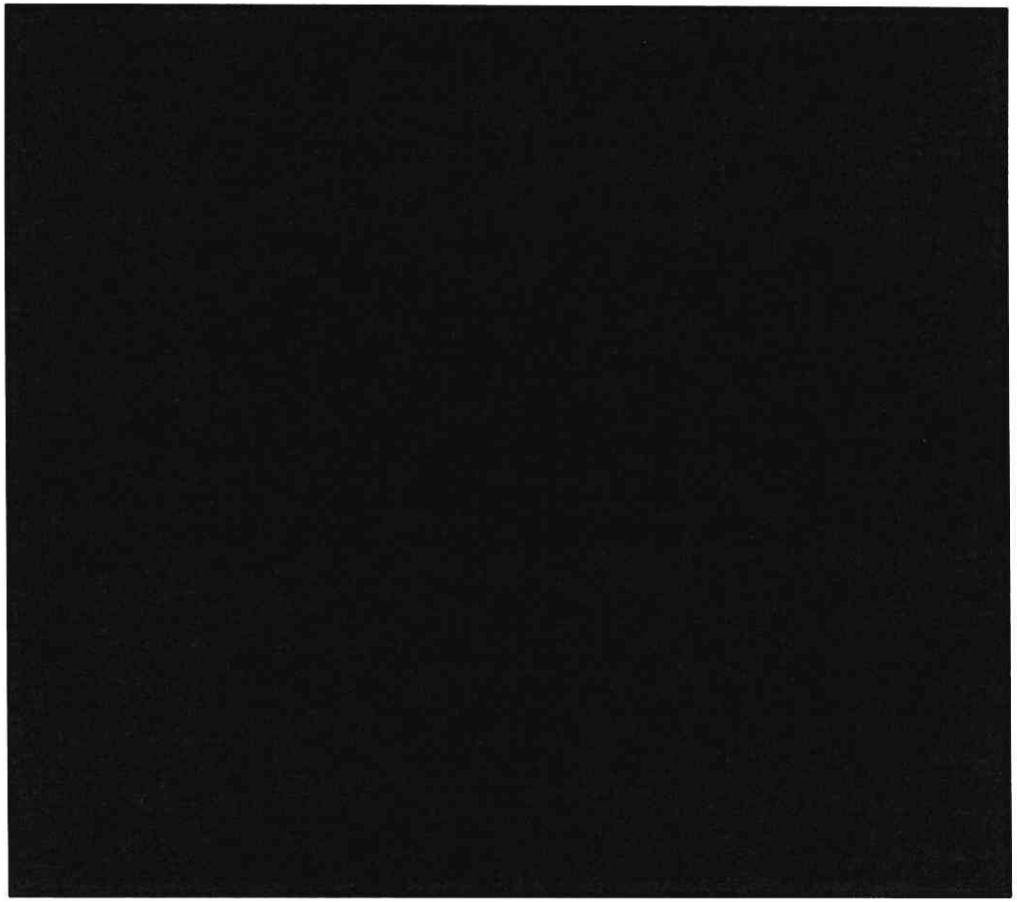
1. Au regard de la non couverture des **exigences** d'une part et des **délais** imposés par le projet EMC2 d'autre part et ce par l'ensemble des offreurs sélectionnés dans le cadre de l'étude, nous recommandons que le projet EMC2 se poursuive sur la **plateforme HDH actuelle**.
2. L'association du HDH à la démarche **démonstrateur cloud de confiance** doit se poursuivre.
3. Les exigences prioritaires du HDH et d'EMC2 ont été intégrées dans **le premier défi :** KMS, HSM et intégration dans les services Cloud, Habilitations : gestion des accès à privilèges, accès conditionnels (cf. annexe 2)
    - Opportunité de communication du lancement des défis du **démonstrateur Cloud** dans le cadre de France 2030 : **Janvier 2024**
    - Première livraison des services du défi 1 en **T4 2024** par les fournisseurs participants
    - Le HDH fait partie des projets pilotes sur plusieurs services et pourra les pré-valider au fil de l'eau pour préparer la migration de sa plateforme
4. Compte tenu du planning du démonstrateur, le HDH devra entamer les travaux de préparation **fin 2024** afin de préparer la mise en œuvre de sa plateforme chez un hébergeur SecNumCloud
5. HDH doit continuer à travailler en priorité sur la migration de sa plateforme sur un **cloud de confiance** pour mise en œuvre en **2025**. Les travaux de construction d'une plateforme spécifique pour EMC2 mettent à risque cet objectif.

V. Annexes

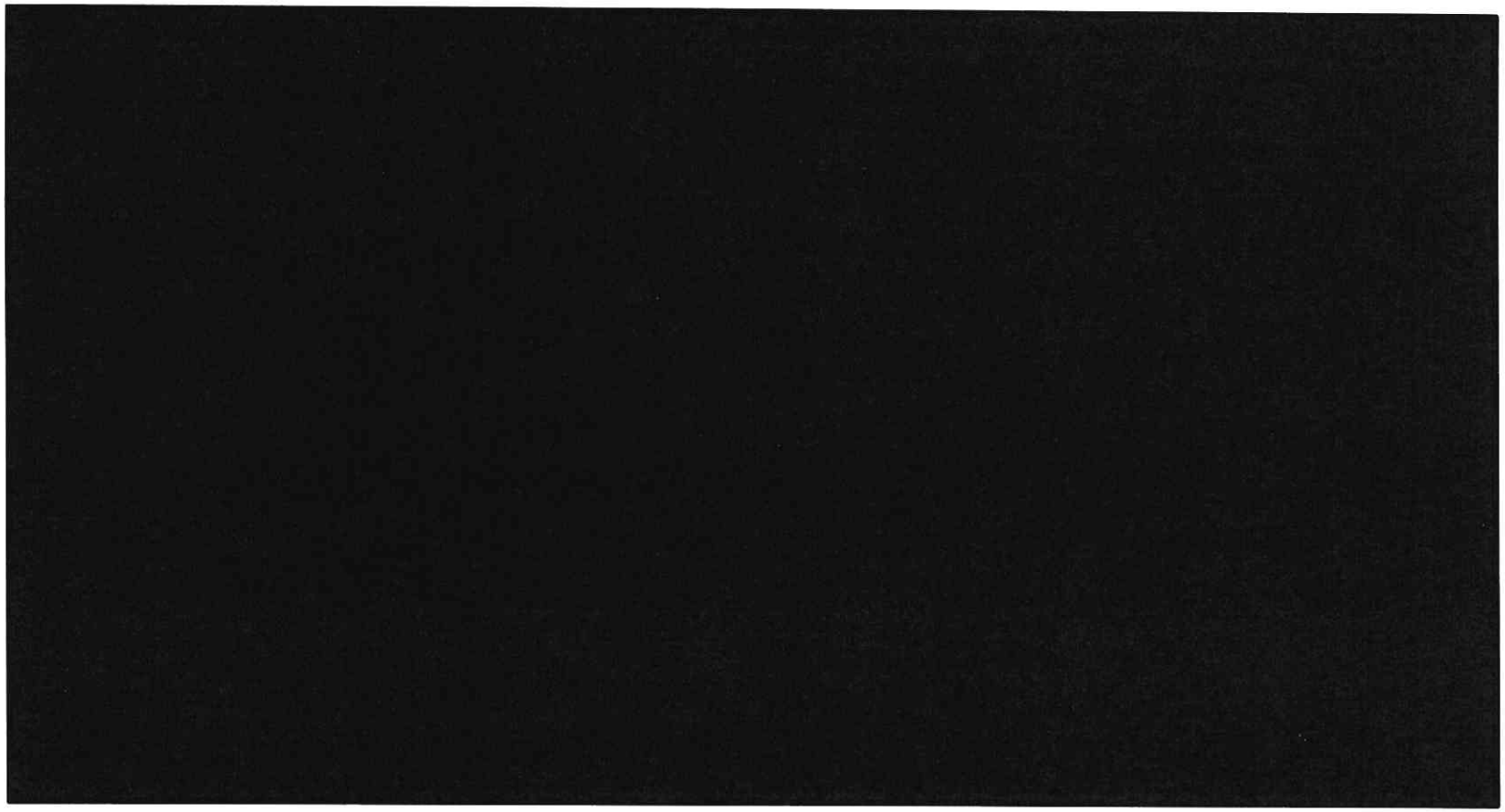
Glossaire

- **EMC2 :** Projet d'entrepôt de données multicentrique pour l'EMA (European Medicines Agency)
- **Bastion** : Serveur ou appareil réseau qui est un point d'entrée dans une zone de confiance forte (réseau interne) depuis une zone de confiance faible (ex internet) et qui est utilisé pour accéder à des ressources sensibles.
- **HSM** (Hardware Security Module): Périphérique matériel dédié à la gestion des clés cryptographiques. Il fournit un environnement sécurisé pour stocker et gérer les clés, ce qui les rend plus difficiles à compromettre.
- **IAM** (Identity and Access Management) : Ensemble de pratiques et de technologies qui permettent de contrôler l'accès aux ressources informatiques.
- **KMS** (Key Management Service) : Service cloud qui fournit des fonctionnalités de gestion des clés pour les applications cloud. Il permet aux organisations de créer, de gérer et de déployer des clés cryptographiques de manière sécurisée.
- **Kubernetes** : Système de gestion de conteneurs open source. Il permet aux organisations de déployer, de gérer et de faire évoluer des applications conteneurisées.
- **LBaaS** (Load Balancing as a Service) : Service cloud qui fournit des fonctionnalités de basculement de charge. Il permet aux organisations de répartir le trafic entre plusieurs serveurs ou services.
- **Micro-segmentation réseau** : Technique de sécurité qui consiste à diviser un réseau en petits segments isolés. Cela permet de limiter la propagation d'une attaque ou d'un incident de sécurité à un seul segment.
- **SIEM** (Security Information and Event Management) : Système de sécurité qui collecte, analyse et corrèle les données de sécurité provenant de différentes sources. Il permet aux organisations de détecter et de réagir aux incidents de sécurité de manière plus efficace.
- **SNC**(SecNumCloud) : La qualification SecNumCloud de l'ANSSI (Agence Nationale de la Sécurité des Systèmes d'Information) s'adresse aux hébergeurs de données prestataires de services cloud et valorise l'aptitude à atteindre un niveau de sécurité élevé.

Annexe 1 : Projet de délibération CNIL du 12 octobre 2023

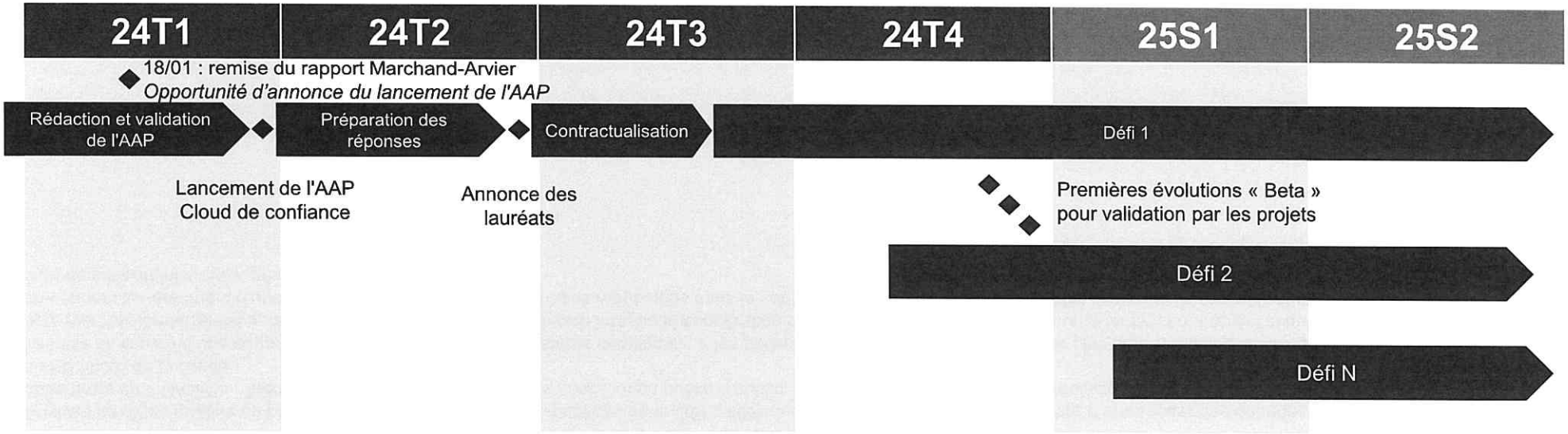


Annexe 1 : Projet de délibération CNIL du 12 octobre 2023



Annexe 2 : Planning Démonstrateur Cloud

Le démonstrateur Cloud de confiance fera l'objet d'un **AAP sur 4 ans lancé au 1er trimestre 2024** par la DGE, le SGPI et la DINUM et animé conjointement avec l'INRIA et l'ANSSI. Les **premières livraisons sont attendues pour le 4ème trimestre 2024**.
Les équipes numériques de l'Administration, dont le HDH, sont associées à la construction de la feuille de route et participeront activement à la validation des services et de l'expérience développeur. Un périmètre a déjà été priorisé pour le lancement d'un premier défi, portant essentiellement sur les capacités de **chiffrement** et **l'amélioration de la gestion des habilitations**.



Le HDH fait partie des projets pilotes sur plusieurs services et les pré-validera pour préparer la migration de sa plateforme

Dans le premier défi :

- Chiffrement : KMS, HSM et intégration dans les services Cloud
- Habilitations : gestion des accès à privilèges, accès conditionnels

Dans les autres défis :

- Containers : orchestration, registres et intégration dans les services Cloud
- Sécurité : puits de logs, SIEM, scan de vulnérabilités et conformité
- Données : catalogue de données, puits de données
- Elasticité du calcul

Annexe 3 : Analyse juridique coopération public-public

Synthèse de l'analyse juridique

Analyse juridique relative au principe d'une coopération public-public entre le GIP « HDH » et le GIP « CASD » pour permettre d'entreposer des données de santé sur la plateforme du « CASD » - Modalités de recours à un accord de coopération public – public, en date du 29 novembre 2023 réalisée par un cabinet d'avocat et revue par la DAJ du Ministère de la Santé.
La mise en œuvre d'une coopération public-public impose plusieurs conditions : « les pouvoirs adjudicateurs doivent réaliser plus de 80 % des activités objet de la coopération hors marché concurrentiel », la coopération public-public est exclue lorsque la coopération entre pouvoirs adjudicateurs a « uniquement pour objet l'acquisition d'une prestation moyennant le versement d'une rémunération »,… qui ne seront pas respectées dans le contexte du projet EMC2 faisant porter un risque de requalification en marché public et à un risque administratif et pénal très important.

Annexe 4 : Retour exigences CNIL/DRESS

Ref exigence	Nom exigence	Description exigence	Retour CNIL	Compte rendu et suites du point CNIL / mission EMC2 du 6/12
R-052	Accès aux informations	L'accès et l'usage des données directement identifiantes doivent être restreints aux finalités listées au point 5.5 et aux seules personnes chargées de la réalisation des opérations nécessaires à l'accomplissement de ces finalités.	**pas applicables au projet EMC2**	Les offreurs de positionnent par rapport cas d'usage, l'exigence est précisé pour motiver et contextualiser l'exigence. Reformuler N/A en pas applicable au projet ECM2 et vérifier que ces exigences ne sont pas incluses dans les taux de couverture des clouders.
R-086	SEC-REI-2	Le cas échéant, et en cas de nécessité dûment justifiée et documentée, le responsable de traitement met en place une procédure opérationnelle sécurisée afin de recontacter des patients pour leur proposer de participer à des recherches. Cette procédure permet, à partir d'une liste de critères médicaux, de sélectionner les identifiants pseudonymes uniques correspondants aux patients visés, puis, en mobilisant la ou les tables de correspondance de l'entrepôt avec ces seuls pseudonymes, de sélectionner les données identifiantes correspondant à ces patients afin de les exporter pour cette seule finalité.	**pas applicables au projet EMC2**	Les offreurs de positionnent par rapport cas d'usage, l'exigence est précisé pour motiver et contextualiser l'exigence. Reformuler N/A en pas applicable au projet ECM2 et vérifier que ces exigences ne sont pas incluses dans les taux de couverture des clouders.
R-087	SEC-REI-3	Le cas échéant, le responsable de traitement met en place une procédure opérationnelle sécurisée afin de ré-identifier des patients en cas d'urgence médicale. Cette procédure permet, en mobilisant la ou les tables de correspondance de l'entrepôt, de sélectionner les données identifiantes des patients concernés à partir de leur numéro pseudonyme unique, et de les exporter pour cette seule finalité.	**pas applicables au projet EMC2**	Les offreurs de positionnent par rapport cas d'usage, l'exigence est précisé pour motiver et contextualiser l'exigence. Reformuler N/A en pas applicable au projet ECM2 et vérifier que ces exigences ne sont pas incluses dans les taux de couverture des clouders.
F-010	Base de données d'un responsable de données-Chaînage avec la base principale via le NIR pseudonymisé	Ingestion de données chaînables directement avec la base principale via le NIR pseudonymisé. Ce cas intervient notamment lorsque la base de données du responsable de données (RD) contient le NIR (ou des traits identifiants qui permettent la reconstruction du NIR par intervention de la CNAV). Cas d'usage : ajouter un objet à un bucket, lister les objets d'un bucket, récupérer un objet depuis un bucket, créer un VPC, supprimer un sous-réseau d'un VPC, supprimer un VPC, accepter un peering entre deux VPC, créer un peering entre deux VPC, refuser un peering entre deux VPC, supprimer un peering entre deux VPC, ajouter un sous-réseau privé à un VPC, instancier un KMS, Créer un cluster Kubernetes, assigner une politique de filtrage à un pare-feu, créer un pare-feu au sein d'un VPC, définir une politique de filtrage (source, destination, protocoles, ports, chiffrement), modifier une politique de filtrage, supprimer une politique de filtrage,	**pas applicables au projet EMC2**	Pas applicable au projet EMC2 Cas d'usage : préciser pas applicable au projet ECM2
R-051	Données à caractère personnel	Dans le cas où des données directement identifiantes, des tables de correspondance, des données génétiques ou des données de suivi de localisation sont versées dans l'entrepôt, celles-ci doivent être stockées séparément des données pseudonymisées, en utilisant les procédés décrits dans les exigences de sécurité SEC-LOG-4 à SEC-LOG-6 Cas d'usage : Activités SEC-LOG-4 et SEC-LOG-6	**pas applicables au projet EMC2**	Pas applicable au projet EMC2 Cas d'usage : préciser pas applicable au projet ECM2 Les données de localisation sont les données de traces GPS.
R-084	SEC-PSE-4	Les documents non structurés ajoutés à l'entrepôt doivent faire l'objet d'une étape de suppression ou de masquage avant leur intégration dans l'entrepôt. Cette étape consiste à supprimer les données identifiantes des patients et des professionnels de santé ou à les remplacer par des termes génériques ou des données fictives. Par exemple, les NIR, nom de naissance, prénom, code postal, ville ou numéro de téléphone seront remplacés par des termes génériques tels que « NIR », « NOM_DE_NAISSANCE », « PRENOM », « CODE_POSTAL », « VILLE » ou « TEL ». Cette exigence s'applique notamment aux documents bureautiques et aux fac-similés d'impression (comme les comptes rendus médicaux et les prescriptions), aux numérisations de documents, à l'imagerie médicale et à toute forme de résultats d'analyse biomédicale. Elle concerne également les commentaires en saisie libres contenus dans les bases de données. L'opération de masquage ou suppression devra s'appliquer au contenu visible des documents (comme les entêtes des courriers et les cartouches des images), aux métadonnées contenues dans ces fichiers (comme le nom de l'opérateur d'imagerie) et aux attributs des fichiers (comme leur nom) Cas d'usage : Couvert par le besoin fonction de l'espace d'ingestion RD	**pas applicables au projet EMC2**	Pas de documents non structurés pour le projet EMC2 Pas applicable au projet EMC2 Cas d'usage : préciser pas applicable au projet ECM2
R-093	Alimentation du SNDS central	Pour l'alimentation du SNDS central, un procédé sécurisé doit être utilisé pour pseudonymiser les données venant des bases sources. Ce procédé doit être basé sur des fonctions cryptographiques robustes répondant aux besoins suivants : être irréversible (impossibilité de disposer d'une transformation inverse permettant de passer d'un pseudonyme à un identifiant initial) ; ne pas générer de collision (deux identifiants initiaux différents donneront deux pseudonymes différents) ; avoir un bon effet d'avalanche (deux identifiants initiaux de valeurs proches donneront deux pseudonymes de valeurs éloignées) ; être une fonction d'agrégation (pour une même transformation, association à un identifiant initial d'un seul et même pseudonyme et association à un seul pseudonyme d'un unique identifiant initial) ; être paramétrable (utilisation possible de différents secrets) ; être identifiable (la fonction utilisée doit être identifiable dans son résultat).Dans le cadre de l'alimentation, la génération des pseudonymes du SNDS central s'opère sur deux niveaux minimum. Les pseudonymes générés pas le gestionnaire du système source cédant les données sont appelés pseudonymes de niveau 1. Ils sont générés au moyen d'une fonction respectant les principes énoncés ci-dessus et alimentent le SNDS central. À la réception de ces jeux de données, le gestionnaire du SNDS central génère de nouveaux pseudonymes (de niveau 2) au moyen d'une fonction respectant les principes énoncés ci-dessus. Les fonctions de pseudonymisation utilisées successivement ne doivent pas avoir les mêmes secrets. Elles ne doivent pas non plus être dérivées les unes des autres, ni être dérivées de fonctions de pseudonymisation déjà existantes. Tous les gestionnaires de systèmes sources alimentent le SNDS central avec un OBJECTIF :Un processus de pseudonymisation doit être assuré au niveau de chaque système afin de protéger la confidentialité des données à caractère personnel. À SAVOIR : Les pseudonymes doivent être différents d'un système à l'autre afin d'éviter qu'une entité ayant des accès sur différents systèmes puisse reconstituer une base complète lui permettant d'identifier indirectement une personne; La confidentialité de la valeur secrète doit être assurée et ses processus de gestion finement décrits; Le renouvellement de pseudonyme doit être assuré en cas de mise à mal du secret de pseudonymisation; A ce jour, le tiers de confiance national et les processus associés ne sont pas encore définis.même pseudonyme de niveau 1. Cas d'usage : Couvert par le besoin fonction de l'espace d'ingestion RD	**pas applicables au projet EMC2**	Pas d'alimentation du SNDS central Pas applicable au projet EMC2 Cas d'usage : préciser pas applicable au projet ECM2

Annexe 4 : Retour exigences CNIL/DRESS

Ref exigence	Nom exigence	Description exigence	Retour CNIL	Brique	Impact sécurité si exigence non considérée pour EMC2	Solution de contournement proposée	Compte rendu et suites du point CNIL / mission EMC2 du 6/12
R-153	HDH-ANSSICNIL-013	Mettre en place un chiffrement des données sur la plateforme Cas d'usage : chiffrer une ressource, activer le chiffrement sur un bucket	Le chiffrement des données est requis mais la granularité du chiffrement n'est pas imposée par les référentiels applicables à EMC2	Sécurité	• Risque de compromission de la donnée lors de l'exposition des buckets à l'extérieur de la plateforme (cas d'usage spécifique EMC2)	N/A	Cette exigence s'impose uniquement dans le cas d'une solution proposant du stockage objet, indispensable pour une architecture cloud. La mission EMC2 précise que le stockage objet n'a pas été imposé au CASD ou à Clinityx, les solutions de stockage HDFS par exemple ont aussi été acceptées. De même pour les solutions IAAS du scénario 1. Dans ce cas, la proposition de solution de stockage impliquera des développements complémentaires pour le HDH pour porter les services applicatifs déjà mis en oeuvre côté HDH.
F-066	HDH-ESS-003-Capacité de résilience au sein d'une région donnée	Besoin Fonctionnel : Garantir la disponiblité des outils Besoin Securité : Disponibilité des outils de monitoring et de suivi des éléments de sécurité Cas d'usage : déployer mon application en haute disponibilité sur au moins deux zones de disponibilité	La haute disponibilité n'est pas imposée par les référentiels applicables à EMC2	Fondation	N/A (besoin fonctionnel)	N/A	Exigence confirmée par l'"équipe projet EMC2. A noter que tous les offreurs savent remplir cette exigence.
F-069	HDH-ESS-007-Centralisation des outils de monitoring	Besoin Securité : Suivi des activités sur la plateforme, traitement des traces Cas d'usage : l'ensemble des opérations soient possibles depuis une API REST, en CLI et depuis la console de management, dans les même conditions de sécurité (certificats, authentification multifacteurs)	La traçabilité est requise mais les API REST ne sont pas imposées par les référentiels applicables à EMC2	Gestion	N/A	Usage CLI ou console de gestion	Les offreurs sont en capacité de fournir les API Rest.
F-070	HDH-ESS-009-Politique de sécurité d'accès aux outils d'administration unique et centralisée	Besoin Securité : - Moindre privilèges sur les éléments du SI - Simplification de la gestion des authentifications Cas d'usage : l'ensemble des opérations soient possibles depuis une API REST, en CLI et depuis la console de management, dans les même conditions de sécurité (certificats, authentification multifacteurs)	Une politique de moindre privilège est requise, la simplification de la gestion des authentifications est conseillée mais non requise et l'API REST n'est pas imposée par les référentiels applicables à EMC2	Gestion	N/A	Usage CLI ou console de gestion	Les offreurs sont en capacité de fournir les API Rest.
R-157	HDH-ANSSICNIL-017	Gestion des clés par un HSM Cas d'usage :Chiffrer et déchiffrer des clé de contenu par le HSM	Une procédure opérationnelle de gestion des clés cryptographiques est requise mais un HSM n'est pas imposé par les référentiels applicables à EMC2	HSM	• **Exposition des clés cryptographiques** : sans un HSM, les clés cryptographiques peuvent être plus vulnérables aux attaques, compromettant l'intégrité et la confidentialité des données. • **Difficulté de conformité au RGPD** : un HSM fournit un environnement sécurisé spécifique à la gestion des clés, renforçant la protection contre les accès non autorisés, ce qui est essentiel pour respecter les exigences du RGPD, particulièrement face aux risques physiques. • **Complexité de gestion** de cycle de vie des clés de chiffrement	1. Utilisation d'un service HSM/KMS souverain tier à la plateforme 2. Gestion manuelle des clés de chiffrement au travers d'un stockage sécurisé 3 types de clés dans le contexte EMC2 : chiffrement des données (repos, objet), chiffrement pour import de clés depuis une solution externe, clés de pseudonymisation	Pour des solutions cloud, l'ANSSI a bloqué des qualifications en raison de l'absence de brique HSM non disponible. La CNIL s'interroge sur la nécessité d'imposer un HSM et se demande si le nombre de clésà gérer pour le projet EMC2 justifie ce type d'équipement alors que peu ou pas d'offreurs sont en capacité de les fournir. Question du nombre de clés pour le projet EMC2 en tenant compte de l'architecture retenue à traiter par équipe EMC2.
T-007	Gestion HSM	Service de gestion relatif à la plateforme HSM Cas d'usage : avoir un point de terminaison privé pour me connecter à mon HSM consulter les logs du HSM relatives à mon rôle modifier le mot de passe d'un autre utilisateur sur le HSM modifier mon mot de passe sur le HSM que le HSM soit qualifié FIPS 140-2 level 3	Une procédure opérationnelle de gestion des clés cryptographiques est requise mais un HSM n'est pas imposé par les référentiels applicables à EMC2	HSM	• **Exposition des clés cryptographiques** : sans un HSM, les clés cryptographiques peuvent être plus vulnérables aux attaques, compromettant l'intégrité et la confidentialité des données. • **Difficulté de conformité au RGPD** : un HSM fournit un environnement sécurisé spécifique à la gestion des clés, renforçant la protection contre les accès non autorisés, ce qui est essentiel pour respecter les exigences du RGPD, particulièrement face aux risques physiques. • **Complexité de gestion** de cycle de vie des clés de chiffrement	1. Utilisation d'un service HSM/KMS souverain tier à la plateforme 2. Gestion manuelle des clés de chiffrement au travers d'un stockage sécurisé 3 types de clés dans le contexte EMC2 : chiffrement des données (repos, objet), chiffrement pour import de clés depuis une solution externe, clés de pseudonymisation	Pour des solutions cloud, l'ANSSI a bloqué des qualifications en raison de l'absence de brique HSM non disponible. La CNIL s'interroge sur la nécessité d'imposer un HSM et se demande si le nombre de clésà gérer pour le projet EMC2 justifie ce type d'équipement alors que peu ou pas d'offreurs sont en capacité de les fournir. Question du nombre de clés pour le projet EMC2 en tenant compte de l'architecture retenue à traiter par équipe EMC2.

Annexe 4 : Retour exigences CNIL/DRESS

Ref exigence	Nom exigence	Description exigence	Retour CNIL	Brique	Impact sécurité si exigence non considérée pour EMC2	Solution de contournement proposée	Compte rendu et suites du point CNIL / mission EMC2 du 6/12
F-068	HDH-ESS-006-Portail de gestion client	Besoin Securité : Mettre en oeuvre la politique de moindre privilèges sur les éléments du SI Cas d'usage : que l'outil d'IAM soit centralisé pour l'ensemble de mes ressources Cloud modifier le rôle d'une ressource Cloud à partir de mon IAM supprimer un rôle d'une ressource Cloud à partir de mon IAM	Une politique de moindre privilège est requise mais l'IAM n'est pas imposé par les référentiels applicables à EMC2	IAM	•**Accès excessif aux ressources et données** : L'absence de granularité dans la gestion des identités peut conduire à des autorisations excessives, où les utilisateurs peuvent avoir accès à des données sensibles au-delà de leur rôle •**Difficulté de gestion des accès** : La gestion des droits d'accès peut devenir complexe, avec un risque accru d'erreurs humaines d'attribution de droits pouvant conduire à des failles de sécurité	1. Création de tenants par projet : gestion segmentée des accès ressources au niveau du projet 2. Utilisation d'un service IAM granulaire et souverain tier à la plateforme	IAM : fourniture de cette brique est discriminante entre offreur Tous les offeurs ne sont pas en capacité de gérer des profils de droits utilisateur, ce qui oblige le Responsable de traitement à créer et paraméter les droits pour chaque compte utilisateur => risque d'erreur et sécurité
F-002	SEGMENTATION-Segmentation des droits d'opération	les droits et accès des opérateurs plateforme sont nativement segmentés, un unique opérateur platefprme ne peut pas réaliser seul une opération sensible de bout en bout. Chacun des accès [privilégiés] des opérateurs sont protégés par l'utilisation d'un bastion. Cas d'usage :assigner une politique de filtrage à un pare-feu créer un pare-feu au sein d'un VPC définir une politique de filtrage (source, destination, protocoles, ports, chiffrement) modifier une politique de filtrage, supprimer une politique de filtrage créer un VPC supprimer un sous-réseau d'un VPC, supprimer un VPC accepter un peering entre deux VPC	Une segmentation réseau est requise mais VPC n'est pas imposé par les référentiels applicables à EMC2	IAM	•**Accès excessif aux ressources et données** : L'absence de granularité dans la gestion des identités peut conduire à des autorisations excessives, où les utilisateurs peuvent avoir accès à des données sensibles au-delà de leur rôle •**Difficulté de gestion des accès** : La gestion des droits d'accès peut devenir complexe, avec un risque accru d'erreurs humaines d'attribution de droits pouvant conduire à des failles de sécurité	1. Création de tenants par projet : gestion segmentée des accès ressources au niveau du projet 2. Utilisation d'un service IAM granulaire et souverain tier à la plateforme	
F-046	SEGMENTATION-Segmentation des environnements	les environnements de développement et de production sont isolés. Un développeur ne peut dès lors effectuer les tâches d'un opérateur Cas d'usage :création des espaces privés virtuels (VPC)pour chaque environnement	Une segmentation des environnements est requise mais VPC n'est pas imposé par les référentiels applicables à EMC2	Kubernetes	Gestion d'un cluster Kubernetes par le HDH et/ou revue de l'architecture logicielle des solutions du HDH.	Installation, configuration et gestion d'un cluster Kubernetes par le HDH	
F-003	SEGMENTATION-Segmentation réseau	la plateforme technologique sera aussi segmentée dans plusieurs sous-réseaux dont le filtrage est assuré par des pare-feux (périmétrique, zone opération, zone service) Cas d'usage : Créer un VPC ajouter un sous-réseau privé à un VPC supprimer un VPC attacher une interface réseau virtuelle à une ressource Cloud créer une interface réseau virtuelle détacher une interface réseau virtuelle d'une ressource Cloud supprimer une interface réseau virtuelle	Une segmentation réseau est requise mais VPC n'est pas imposé par les référentiels applicables à EMC2	Réseau	• Risque d'accès au stockage depuis l'ensemble des bulles sécurisées au sein de l'entrepôt de données • Détection plus complexe et moins rapide des comportements anormaux (supervision plus large)	• Mise en place de firewall pour chaque service d'infrastructure (stockage, VDI, bastion, serveurs de logs, conteneurs, etc.) à 100 firewalls environ pour EMC2	Les solutions de VPC ou de VLAN sont acceptées mais pour mettre en place un niveau de filtrage fin des accès la solution VPC évite de multiplier le nombre de firewall. Remarque générale : le choix d'une solution cloud permettant de bénéficier de fonctions de scalabilité et d'élasticité de la puissance machine est un choix projet qui a des conséquences en termes d'exigences technique sur le projet EMC2 et explique du point de vue de la misison un niveau d'exigence technique supérieure aux exigences du référentiel EDS.

Annexe 5 : Calendrier des ateliers de la mission

	S1			S2					S3					S4					S5					S6					S7					S8		
Type d'atelier	**25/10**	**26/10**	**27/10**	**30/10**	**31/10**	**1/11**	**2/11**	**3/11**	**6/11**	**7/11**	**8/11**	**9/11**	**10/11**	**13/11**	**14/11**	**15/11**	**16/11**	**17/11**	**20/11**	**21/11**	**22/11**	**23/11**	**24/11**	**27/11**	**28/11**	**29/11**	**30/11**	**1/12**	**4/12**	**5/12**	**6/12**	**7/12**	**8/12**	**11/12**	**12/12**	**13/12**
Suivis internes									2				2																							
Atelier de travail HDH																																				
Atelier de travail ANSSI																																				
Atelier de travail CNIL																																				
Atelier offreur																2					2				2												
COPIL																																				

★ Contact Clouders (02/11)

★ Présentation catalogue V0 (10/11)

★ Présentation catalogue V1 (16/11)

Réponses fournisseurs consolidées (01/12)

Livraison du rapport VF (13/12)