Abilian
CS LAB

OPEN INTEROP · S4

Réversibilité du stockage objet : constat et propositions

Stefane Fermigier (Abilian SAS) · 11 novembre 2026 · DRAFT

Version 0.4
Licence CC BY 4.0

Cette note est la note de fond du projet S4 (*Sovereign S3*), le chantier du GT Cloud d'Open Interop sur la réversibilité du stockage objet compatible S3. Elle expose pourquoi le stockage objet sert de cas pilote, les limites de la « compatibilité S3 », le cadre réglementaire et normatif, puis les propositions du projet : un profil versionné en niveaux, un manifeste d'export, des contrats comportementaux, une scorecard et un kit de test de conformité (CTK, *Compliance Test Kit*). Elle intègre les décisions de l'atelier Open Interop du 9 octobre 2026 (Track 6, « Connecting the Clouds », axe Cloud IaaS/PaaS) et les premiers résultats du CTK S4.

Les autres notes de la série détaillent chaque élément. Ce sont la synthèse, les contrats comportementaux, la scorecard, les scénarios de test, l'implémentation du CTK S4, la note Participer aux essais croisés et le guide Par où commencer. Toutes sont des versions de travail, soumises à la relecture du groupe de travail Cloud.

Synthèse

1. **Le stockage objet est le bon cas pilote.** C'est le seul service managé doté d'une interface de fait quasi universelle, l'API S3. Si la réversibilité ne peut pas être rendue opposable ici, elle ne le sera pour aucun autre service. La compatibilité S3 ne suffit pourtant pas : les écarts se logent dans les comportements, les métadonnées, les limites de débit et les fonctions annexes.
2. **L'API S3 n'est pas un standard ouvert** au sens de l'EIFv1. AWS en contrôle seul la spécification et la fait évoluer sans concertation. Une exigence opposable ne peut donc pas renvoyer simplement à « l'API S3 » : elle doit citer un **profil S4** ouvert, versionné et gouverné collectivement, dont le niveau Core constitue le périmètre de la première version. S4 formalise le standard de fait qu'est devenue l'API S3, pour les fonctions utiles à la réversibilité : une version datée, indépendante des évolutions décidées par AWS, avec des comportements vérifiables.
3. **Certaines métadonnées ne transitent pas par l'API** (date de dernière modification, identifiants de version, ETag multipartie, rétention). La réversibilité exige un **manifeste d'export** qui accompagne les données et conserve leurs valeurs d'origine.
4. **La conformité repose sur des critères obligatoires, sans compensation.** Dans une première étape, cinq critères doivent être satisfaits à 100 % : complétude, intégrité, métadonnées obliga-

toires, droits d'accès et portabilité fonctionnelle. Durée, débit, reprise, coût et effort sont publiés à part, avec leurs conditions. Aucun score pondéré ne fonde un verdict.

5. **Le CTK S4 fournit une preuve technique reproductible.** Il met en œuvre les scénarios CTK-01 à CTK-05. Ses premiers essais révèlent des écarts qu'aucune déclaration « compatible S3 » ne signale (§4.6). Son usage dans un contrat demande une rédaction propre, attendue en décembre 2026.
6. **Les cinq niveaux d'interopérabilité sont couverts.** Les niveaux organisationnel (assistance, délais) et juridique (effacement, preuve, contrat) sont en grande partie cadrés par le chapitre VI du Data Act et relèvent du contrat et de l'audit autant que du test.

1. Pourquoi le stockage objet

1.1 La question posée

La question de départ est celle de l'acheteur : *si un acheteur exige demain la réversibilité d'un service cloud, comment vérifiera-t-il qu'elle est satisfaite ?* La réponse attendue est une chaîne de preuve : contrat comportemental, métrique, test reproductible, seuil, verdict de conformité.

L'atelier du 9 octobre 2026 l'a formulée ainsi. L'acheteur doit pouvoir exécuter, ou faire exécuter, un test versionné sur un jeu de données connu, récupérer un rapport vérifiable et constater que ses données sont utilisables chez un autre fournisseur. Une déclaration « compatible S3 » doit préciser les fonctions prises en charge et s'accompagner d'une preuve sur le périmètre revendiqué. Cette approche applique la définition de l'interopérabilité opposable : exiger un résultat mesurable et vérifier qu'il est atteint, sans imposer de technologie.

1.2 Raisons du choix

- **Le stockage objet concentre l'essentiel des volumes** : sauvegardes, archives, lacs de données, contenus applicatifs, artefacts. C'est souvent le plus gros volume à rapatrier lors d'une sortie.
- **Une interface de fait existe.** L'API S3 est implémentée par la quasi-totalité des fournisseurs européens (OVHcloud, Scaleway, Outscale, Cloud Temple, Clever Cloud, IONOS, etc.) et par plusieurs logiciels libres (Ceph RGW, Garage, SeaweedFS, MinIO…).
- **Le cas paraît simple.** Si la compatibilité S3 suffisait, la réversibilité du stockage objet serait un problème résolu ; la section 2 montre qu'elle ne l'est pas.
- **La méthode se généralise.** Profil d'interface, format d'export, contrats comportementaux, CTK et scorecard pourront ensuite s'appliquer aux bases de données managées, aux files de messages, aux secrets et aux journaux.

1.3 Place dans le plan d'action CSF Cloud et dans Open Interop

Le projet S4 met en œuvre plusieurs actions du plan d'action de la filière cloud, résumé dans la fiche cloud. Les contrats comportementaux relèvent du pilier 1 (action 1.2, dont l'exemple est l'export massif d'un stockage objet) ; l'indice d'interopérabilité et le CTK, du pilier 2 ; les clauses de marchés publics, du pilier 3 ; les TechSprints et l'autorité technique, des piliers 4 et 5. Il donne au chantier CL4 (indicateur d'interopérabilité cloud) une orientation par le résultat : sortir effectivement ses données, sans perte, dans un délai et à un coût connus. Un score fondé sur la liste des standards déclarés ne mesurerait qu'une intention.

2. Les limites de « compatible S3 »

La compatibilité S3 tient pour l'usage courant : écrire, lire et lister des objets dans un bucket non versionné. Une sortie complète révèle des écarts de plusieurs natures.

2.1 Écarts de couverture fonctionnelle

Chaque fournisseur implémente un sous-ensemble différent de l'API. Les fonctions le plus souvent absentes ou partielles contiennent aussi l'état à préserver :

- versionnage et marqueurs de suppression (*delete markers*) ;
- verrouillage d'objets (*Object Lock*, modes *governance* et *compliance*, *legal hold*) ;
- étiquettes d'objets (*tagging*) ;
- règles de cycle de vie, CORS, chiffrement par défaut, notifications, réplication ;
- classes de stockage, notamment les classes froides qui exigent une restauration préalable ;
- requêtes conditionnelles (`If-Match`, `If-None-Match` en écriture, ajoutées par AWS en 2024) ;
- sommes de contrôle additionnelles (CRC32, CRC32C, SHA-1, SHA-256, CRC64NVME).

Les essais du CTK S4 en donnent un exemple : Garage v1.0.1 répond `NotImplemented` aux opérations sur les étiquettes d'objets. Le service reste conforme au niveau Core, qui ne les exige pas. Il devient non conforme dès qu'il revendique les étiquettes.

2.2 Écarts de comportement à interface identique

À opérations identiques, les comportements divergent :

- **Pagination et ordre de listage** : ordre lexicographique UTF-8 des clés, taille de page, stabilité du jeton de continuation sur de très grands buckets, cohérence du listage après écriture.
- **Erreurs et limitation de débit** : codes HTTP, codes d'erreur S3 (`SlowDown`, `ServiceUnavailable`, `InternalError`), en-têtes `Retry-After`, seuils de limitation par préfixe ou par compte, rarement documentés.
- **Clés d'objets difficiles** : caractères spéciaux, Unicode non normalisé (NFC ou NFD), clés très longues (jusqu'à 1 024 octets chez AWS), barres obliques consécutives, segments `.` et `..`, objets « répertoires » de taille nulle terminés par `/`.
- **Adressage** : style *virtual-hosted* ou *path-style*, noms de buckets globaux (AWS) ou propres à une région ou un projet (la plupart des autres), règles de nommage différentes.
- **Envoi multipartie** : taille minimale des parties, nombre maximal de parties, taille maximale d'objet, traitement des envois incomplets (nettoyage, facturation).

Le jeu de données du CTK S4 contient un objet multipartie dont la dernière partie fait un octet (5 Mio + 1). Sur Ceph RGW v19.2.7 et v20.2.4, avec le moteur de stockage SQLite (`dbstore`) employé pour les essais, la lecture de cet objet ne se termine jamais et le serveur reste à 100 % de CPU. Un cas limite que rien ne signale dans la documentation suffit à bloquer une migration.

2.3 Métadonnées non reproductibles par l'API

Plusieurs propriétés ne peuvent pas être recréées à l'identique chez le fournisseur cible, même par un outil parfait, parce que l'API les fait attribuer par le serveur :

Propriété	Cause de la perte	Conséquence
`Last-Modified`	Fixée par le serveur au moment de l'écriture	Perte de l'historique temporel ; effets sur les règles de cycle de vie, les audits et la conservation légale

Propriété	Cause de la perte	Conséquence
Identifiants de version (`VersionId`)	Générés par le serveur	Les références externes vers une version donnée sont cassées
Ordre et dates des versions, marqueurs de suppression	Recréables seulement par rejeu chronologique ; les dates restent celles du rejeu	Historique faussé
`ETag` des objets multiparties	Dépend du découpage en parties (MD5 des MD5 des parties, suffixé du nombre de parties) ; n'est pas un MD5 en cas de chiffrement SSE-KMS	Inutilisable pour vérifier l'intégrité de bout en bout
Rétention Object Lock	Doit être réappliquée, avec une durée restante recalculée	Risque de non-conformité réglementaire (archivage légal, WORM)
Propriétaire, ACL, politiques de bucket	Liés au modèle IAM du fournisseur	Doivent être traduits, pas copiés

La réversibilité ne peut donc pas reposer sur la seule recopie par l'API. Il faut un artefact qui transporte les informations que l'API ne transporte pas : le manifeste d'export (§4.2).

2.4 Écarts liés à l'environnement du service

- **Classes froides.** Un objet archivé doit être restauré avant lecture, avec des délais de plusieurs heures et parfois des coûts propres. Pour une sortie, ce délai domine souvent la durée totale.
- **Chiffrement.** SSE-S3 et SSE-KMS sont transparents à la lecture si les droits sur la clé KMS sont accordés ; SSE-C exige de fournir la clé à chaque requête ; un chiffrement côté client suppose que le client a conservé ses clés. L'export des clés KMS elles-mêmes (BYOK, HYOK) déborde du stockage objet, mais conditionne la réversibilité.
- **Réseau et débit de sortie.** Le débit soutenable dépend des limites du fournisseur (bande passante sortante, limitation de requêtes) autant que du client. Ordres de grandeur : 1 Po à 10 Gbit/s demande environ 9,3 jours, comme 100 To à 1 Gbit/s. Pour les très gros volumes, l'existence d'un service d'export assisté (transfert physique, liaison dédiée temporaire) devient déterminante.

2.5 Une interface contrôlée par un seul acteur : l'incident des sommes de contrôle de 2025

En janvier 2025, les SDK AWS (boto3 à partir de la version 1.36, AWS CLI, SDK Java, etc.) ont activé par défaut de nouvelles protections d'intégrité : calcul systématique de sommes de contrôle CRC à l'envoi et validation à la réception. De nombreux services et logiciels « compatibles S3 » ont alors cessé de fonctionner avec ces versions, jusqu'à ce que les utilisateurs modifient leur configuration ou que les fournisseurs adaptent leurs implémentations.

Le CTK S4 reproduit l'incident contre Garage : un client boto3 dans sa configuration par défaut ne parvient pas à relire un objet multipartie qu'il vient d'envoyer. Le client de référence du CTK limite donc les sommes de contrôle aux cas où l'API les exige et vérifie l'intégrité par sa propre empreinte SHA-256.

L'épisode montre deux choses. La compatibilité S3 est une cible mouvante, définie par un seul acteur. Les outils clients employés pour sortir d'un cloud ne sont pas neutres : une mise à jour de SDK peut faire échouer une migration sans que le service ait changé. Un CTK doit donc fixer la version du profil et la configuration des outils de référence, puis les consigner dans chaque rapport.

3. Cadre réglementaire et normatif

Les briques transverses souvent citées (OIDC, OAuth, OpenAPI 3.1, TLS 1.3) relèvent du socle transverse et concernent peu ce cas d'usage. L'API S3 s'authentifie par signature de requête (SigV4) et OpenAPI ne décrit pas les comportements en cause. Deux plans sont à distinguer : l'identité de l'outil qui orchestre la migration (elle peut reposer sur OIDC) et la signature des requêtes S3 elles-mêmes. Pour une exigence formelle, OAuth 2.1 reste un projet de l'IETF : il faut référencer OAuth 2.0 et ses extensions de sécurité publiées, dont la RFC 9700.

3.1 Data Act (règlement (UE) 2023/2854), chapitre VI

Le chapitre VI (« Changement de service de traitement de données »), applicable depuis le 12 septembre 2025, est le texte le plus directement opposable. Il fournit déjà une partie du contrat de réversibilité :

Disposition	Contenu	Usage pour S4
Art. 23	Obligation de lever les obstacles (précommerciaux, commerciaux, techniques, contractuels, organisationnels) au changement de fournisseur	Fondement général de l'opposabilité
Art. 25	Clauses contractuelles obligatoires : préavis maximal de 2 mois, période de transition de 30 jours (jusqu'à 7 mois si elle est techniquement impossible), assistance, période de récupération d'au moins 30 jours, effacement complet des données exportables à l'issue, spécification exhaustive des catégories de données exportables (y compris les métadonnées)	Contrats organisationnels et juridiques (famille ORG)
Art. 26	Obligations d'information : procédures, méthodes et formats de portage, limitations techniques connues, registre en ligne des structures et formats de données et des spécifications ouvertes utilisées	Base légale de la déclaration du fournisseur ; le manifeste d'export peut s'y rattacher
Art. 29	Suppression progressive des frais de changement : frais limités aux coûts pendant la période transitoire, suppression totale à partir du 12 janvier 2027	Le critère « coût » est en grande partie tranché par la loi
Art. 30	Aspects techniques : équivalence fonctionnelle pour l'IaaS ; interfaces ouvertes et gratuites, compatibilité avec les spécifications communes et normes harmonisées pour les autres services ; export dans un format structuré, couramment utilisé et lisible par machine	Le stockage objet se situe à la frontière IaaS/PaaS ; l'équivalence fonctionnelle reste à rendre opérationnelle
Art. 35	Spécifications d'interopérabilité ouvertes et normes harmonisées, couvrant notamment les aspects syntaxiques, sémantiques et comportementaux ; demandes de normalisation ; spécifications communes adoptées par la Commission ; référentiel central de l'Union	Débouché européen du profil S4 et du CTK ; fondement des contrats comportementaux

Une partie de l'opposabilité existe donc déjà. Le travail consiste à rendre ces obligations vérifiables techniquement : définir l'équivalence fonctionnelle d'un bucket, l'export complet, la preuve de l'effacement.

L'article 35 offre un débouché : un profil ouvert et un CTK produits par une communauté européenne peuvent alimenter une demande de normalisation ou des spécifications communes.

Les trois grands hébergeurs américains ont annoncé en 2024 la gratuité du trafic sortant en cas de sortie définitive, sous conditions (demande préalable, délai, fermeture du compte). Le critère de coût doit donc intégrer les conditions d'accès à cette gratuité en plus du tarif.

3.2 Cadre français

- **Loi SREN** (loi n° 2024-449 du 21 mai 2024) : encadrement des crédits cloud, limitation des frais de transfert, exigences d'interopérabilité et de portabilité dont l'ARCEP précise les modalités (consultation d'octobre 2024, à laquelle ont notamment répondu le CNLL et Euclidia, puis recommandation d'octobre 2025). Les travaux S4 peuvent alimenter ces spécifications. Les recommandations de transparence de l'ARCEP ne confèrent toutefois pas d'elles-mêmes une portée juridique au CTK.
- **Circulaire du Premier ministre n° 6519/SG du 5 février 2026** sur la commande publique numérique : la réversibilité y est « un élément-clé dans le choix des solutions » qui « doit être prise en compte dans les contrats », sans référentiel technique pour l'évaluer. La scorecard et le CTK peuvent combler ce manque.
- **SecNumCloud** (ANSSI) : le référentiel comporte des exigences de réversibilité et de fin de contrat, avec lesquelles le CTK doit s'articuler pour éviter des exigences concurrentes.
- **RGI** : la note sur le RGI précise comment les profils Open Interop peuvent être mis à la disposition de l'État.

3.3 Normes, codes de conduite et formats existants

Référence	Apport
ISO/IEC 19941:2017 (interopérabilité et portabilité du cloud)	Vocabulaire ; sa facette comportementale correspond aux contrats comportementaux
ISO/IEC 19086 (accords de niveau de service cloud)	Modèle de définition des métriques, pour formaliser la scorecard de façon compatible avec les SLA
SWIPO (codes de conduite IaaS et SaaS, 2020)	Liste d'informations de portage ; déclaratif, utile comme liste de contrôle
SNIA CDMI (ISO/IEC 17826)	Seule norme formelle d'interface de stockage cloud, peu adoptée ; modèle de métadonnées et d'export
SNIA Cloud Object Storage	Plugfests et outils de test : seul effort multifournisseur actif sur la compatibilité S3 (§3.4)
OpenStack Swift, **CNCF COSI**	API ouverte minoritaire ; normalisation du seul plan de contrôle depuis Kubernetes
BagIt (RFC 8493)	Modèle d'empaquetage de l'export (inventaire et empreintes)
S3 Inventory (AWS)	Inventaire propriétaire des objets ; précédent pour le manifeste

L'état de l'art situe ces références parmi les standards du cloud.

3.4 État de la normalisation de l'API S3

Aucun organisme de normalisation (IETF, ISO/IEC, ETSI, CEN) ne mène aujourd'hui de travaux sur l'API S3 ; aucun Internet-Draft n'est en cours. AWS n'a jamais annoncé vouloir confier l'interface à une instance neutre.

L'effort le plus proche est celui de SNIA, dont les Plugfests de 2024 et 2025 ont relevé ambiguïtés, écarts de contrôle d'accès, en-têtes erronés et appels non pris en charge. Son groupe de travail prépare des suites de tests libres, mais ne rédige ni spécification ni profil de conformité ; ses Plugfests se tiennent sous accord de confidentialité, un cadre adapté à la correction de bogues, inadapté à l'opposabilité.

Un article récent (A. Brandes, *The Substitution Escrow Threshold*, arXiv, août 2026) propose cinq conditions pour qu'une revendication de compatibilité réduise réellement le risque de l'acheteur, parmi lesquelles la conformité exécutable, l'indépendance vis-à-vis du détenteur de la spécification et la réversibilité de l'état. Il analyse S3 parmi cinq cas. Ce cadre recoupe le profil S4 (§4.1) et la démarche de vérification (§4.5).

4. Propositions

4.1 Un profil S4 ouvert, versionné en niveaux

Renvoyer à « l'API S3 » dans une exigence contractuelle laisse un flou : ni le sous-ensemble ni la version ne sont précisés. Le texte de référence dépend en outre d'un tiers (§2.5). S4 formalise le standard de fait qu'est devenue l'API S3, pour les fonctions utiles à la réversibilité : une version datée, indépendante des évolutions décidées par AWS, avec des comportements vérifiables. Le projet propose pour cela un **profil S4**, profil d'application au sens de la doctrine :

- maintenu par une instance neutre selon un processus ouvert et validé par l'autorité technique que prévoit la doctrine proposée ;
- versionné, daté et figé (par exemple « S4 2026.1 »), indépendamment des évolutions décidées par AWS ;
- décrivant les opérations et les comportements attendus (pagination, erreurs, limitation, clés, envoi multipartie), qui font l'objet des contrats comportementaux ;
- structuré en niveaux, pour que chaque fournisseur déclare son périmètre exact :

Niveau	Contenu	Usage typique
Core	Buckets, objets, métadonnées système et utilisateur, listage paginé, envoi multipartie, sommes de contrôle, erreurs et limitation de débit	Exigible de tout service se déclarant compatible S3 ; périmètre de la première version
Versioned	Versionnage, marqueurs de suppression, listage des versions	Sauvegardes, données à historique
Governance	Object Lock (*governance*, *compliance*), *legal hold*, étiquettes	Archivage légal, conformité
Extended	Cycle de vie, classes de stockage, CORS, notifications, chiffrement par défaut	Fonctions avancées, traitées au cas par cas

L'atelier du 9 octobre 2026 a retenu le principe : un sous-ensemble documenté de l'API S3, avec déclaration des fonctions prises en charge et preuve sur ce périmètre. Le niveau Core constitue la première version du profil ; les niveaux Versioned, Governance et Extended forment les « périmètres avancés », avec la cohérence forte, la migration pendant les écritures et les dépendances aux services de gestion de clés. Le

CTK S4 couvre aujourd'hui le niveau Core et les étiquettes. La réunion du groupe de travail du 2 novembre 2026 a rappelé que la compatibilité S3 ne suffit pas à l'objectif et demandé de tester une couverture identifiée et datée, chaque fournisseur explicitant les fonctions prises en charge et leurs limites.

Articulation avec SNIA. Le profil reprendrait les suites de tests de SNIA dès leur publication. Open Interop apporterait les éléments absents de ces travaux : un profil publié et versionné, des résultats de CTK publics, le manifeste d'export et un débouché réglementaire au titre de l'article 35 du Data Act (demande de normalisation au CEN-CENELEC, ou ISO/IEC JTC 1/SC 38 via l'AFNOR).

Nom. AWS détient les marques « S3 » et « Amazon S3 ». Le profil décrit l'interface sans s'approprier la marque ; la compatibilité du nom « S4 » avec cette contrainte reste à vérifier.

4.2 Un manifeste d'export

Une partie de l'état d'un bucket ne transite pas par l'API (§2.3). L'article 26 du Data Act exige déjà une information structurée sur les données exportables. Le **manifeste d'export** accompagne les données. Produit par le fournisseur source ou par l'outil d'export, il permet :

- de vérifier la complétude (tous les objets et, aux niveaux avancés, toutes les versions) ;
- de vérifier l'intégrité de bout en bout par une empreinte de l'objet complet, indépendante du découpage multipartie : SHA-256 ;
- de transporter les valeurs d'origine des métadonnées non reproductibles (dates, identifiants de version, rétention, propriétaire), conservées à côté des données chez la cible ;
- de décrire la configuration du bucket (versionnage, cycle de vie, CORS, Object Lock, politique d'accès sous forme documentaire).

Le format retenu est un fichier JSON Lines, une ligne par objet ou par version, décrit par un JSON Schema ; ses champs sont détaillés en annexe de la note sur les scénarios. L'empaquetage peut s'inspirer de BagIt.

L'atelier a fixé deux règles d'interprétation. L'ETag n'est pas une preuve d'intégrité : le CTK calcule sa propre empreinte. Le manifeste préserve les valeurs d'origine des identifiants de version et des dates système sans prétendre conserver leur fonctionnement natif ; la cible ne peut pas, par exemple, servir une version par son ancien identifiant.

Le CTK S4 met en œuvre une première version du schéma, limitée au niveau Core : il produit trois manifestes par essai (référence, source, cible), les compare clé par clé et valide les manifestes fournis par d'autres outils.

4.3 Des contrats comportementaux

Les contrats comportementaux décrivent les garanties du service au-delà de la forme de son interface. La doctrine proposée les nomme « contrats sémantiques ». Dans le corpus Open Interop, « sémantique » désigne aussi le troisième niveau d'interopérabilité (la signification des données). Le terme « contrat comportemental », repris de la facette comportementale de l'ISO/IEC 19941, évite la confusion. Les deux termes désignent la même chose.

Les contrats sont regroupés en familles : export (EXP), métadonnées (MET), intégrité (INT), performance (PERF), comportements d'interface (API), reconstruction (REC), sécurité et droits (SEC), reprise après incident (REP) et exigences organisationnelles et juridiques (ORG). Chacun reçoit un niveau de maturité : complétude et intégrité sont prêtes à être spécifiées, la migration pendant les écritures n'est pas mûre.

Trois contrats précisent les exigences les plus récentes. SEC-01 demande que les droits reconstruits chez la cible soient équivalents, comparés par leur effet sur les accès, sans copier les politiques d'un modèle

IAM à l'autre. REP-01 demande qu'une migration interrompue reprenne sans perte ni duplication. REC-04 définit la portabilité fonctionnelle : une application doit fonctionner sur la cible après révocation de ses accès à la source.

4.4 Une scorecard sans compensation

Agréger intégrité, débit et coût en un score unique reviendrait à accepter qu'un excellent débit compense une perte de données. La scorecard reprend la règle de la doctrine proposée (§5.2) : la conformité exige que tous les critères obligatoires du périmètre revendiqué soient satisfaits, sans compensation ; aucune moyenne pondérée ne fonde un verdict. Elle comporte trois blocs :

1. **Critères obligatoires**, en tout ou rien. La première étape en retient cinq, ceux de l'atelier : complétude, intégrité, métadonnées obligatoires, droits d'accès, portabilité fonctionnelle. Les cinq autres critères du bloc A (configuration, comportement d'interface, réimport, indépendance, automatisation) s'ajouteront dans une étape suivante. Un objet manquant, un contenu corrompu ou un droit interdit devenu autorisé suffit à exclure la conformité sur le profil testé.
2. **Mesures publiées** (B1 à B12) : débit, limites de requêtes, délais de listage et de restauration, reprise, durée totale, débit utile et ressources consommées. Elles sont publiées avec leurs conditions de mesure ; l'acheteur fixe ses seuils selon son besoin. Un seuil universel de débit serait prématuré.
3. **Contexte** (C1 à C8) : coût de sortie et conditions de gratuité, part automatisable du processus, coût et effort observés, dépendances résiduelles, niveau de profil déclaré, délais contractuels, qualité de la documentation.

La restitution prend la forme d'une fiche de profil par fournisseur. D'éventuels paliers se définiraient par les niveaux de profil et les critères obligatoires satisfaits, jamais par une moyenne.

4.5 L'opposabilité par la vérification

Pour les critères obligatoires, le seuil est fixé par le profil et il est absolu : zéro perte, zéro corruption. Pour les grandeurs de performance, un seuil universel serait soit trop bas pour être utile, soit hors de portée des petits fournisseurs ; il dépend en outre du volume, du réseau et du client. Le profil fixe donc la méthode de mesure. Le fournisseur déclare ses valeurs, conformément à l'article 26 du Data Act. Le CTK mesure dans des conditions consignées. L'acheteur fixe ses seuils dans son cahier des charges. La comparaison entre valeurs déclarées et mesurées, avec une tolérance à définir, est l'option envisagée pour le scénario CTK-06.

Chaque critère reçoit l'un des quatre verdicts du profil S4 : conforme, non conforme, non applicable (hors du périmètre revendiqué), non concluant (l'environnement de test a failli). Une fonction revendiquée mais non prise en charge donne « non conforme ». La preuve s'établit à trois niveaux (doctrine, §5.3) : autoévaluation par le fournisseur, vérification par un acheteur ou un tiers qui rejoue l'essai, essai croisé organisé par Open Interop. Seul ce dernier ouvre droit au badge délivré par l'autorité technique, qui tranche aussi les différends.

Le rapport du CTK fournit une preuve technique, assez complète pour qu'un tiers rejoue l'essai. Une exigence devient exploitable dans un contrat quand son périmètre, sa mesure, ses conditions de test et ses conséquences en cas d'écart sont explicites : l'usage contractuel du CTK demande donc une rédaction propre, celle de l'annexe contractuelle proposée pour décembre 2026 (§6).

Certaines exigences échappent au test. L'effacement des données chez la source ne peut pas être vérifié de l'extérieur : il relève d'une déclaration auditable (attestation, journal, contrôle par un tiers). Ces exigences sont opposables par le contrat et l'audit.

4.6 Le CTK S4

Les scénarios partent d'un cas de référence fixé par l'atelier. Un client migre un stockage objet d'un fournisseur A vers un fournisseur B avec un outil indépendant, sur une source figée. Il part d'un jeu de données synthétique reproductible : environ 10 Go et 10 000 objets pour le profil fonctionnel, 1 To et 100 000 objets pour le profil de charge. Il doit retrouver les contenus et propriétés convenus, puis utiliser ses données sans dépendre de A, y compris après révocation de ses accès à A. Six scénarios en découlent : export complet (CTK-01), intégrité avec corruptions volontaires (CTK-02), métadonnées et droits (CTK-03), incident et reprise (CTK-04), exploitation sur B (CTK-05), performance de la migration (CTK-06, à définir). Quatre scénarios complémentaires (CTK-X1 à CTK-X4) couvrent la grande cardinalité, le manifeste, les versions et les classes froides.

Le CTK S4, publié sous licence Apache 2.0 (https://git.sr.ht/~sfermigier/s4-ctk), met en œuvre CTK-01 à CTK-05 avec les quatre verdicts. Il sert d'outil de transfert indépendant et se teste contre sept implémentations libres de l'API S3, dont Garage (Deuxfleurs, adapté au niveau Core) et Ceph RGW (couverture plus large, adaptée aux niveaux Versioned et Governance), ainsi que MinIO, SeaweedFS, Versity Gateway, RustFS et leBLOB. Ses premiers essais, au volume d'essai rapide, illustrent les écarts décrits à la section 2 :

- Garage est conforme au niveau Core sur les cinq scénarios, sauf sur les lectures conditionnelles (`If-Match`), qu'il ignore ; il est non conforme dès qu'il revendique les étiquettes (§2.1) ;
- Ceph RGW, sur son moteur de stockage autonome, ne termine jamais la lecture d'un objet multipartie à dernière partie d'un octet (§2.2) ;
- un client boto3 par défaut échoue contre Garage : l'incident de 2025 est reproduit (§2.5).

Des outils libres existants complètent le kit : ceph/s3-tests, la suite de compatibilité S3 la plus complète, pour tester le profil opération par opération ; rclone, l'outil de migration entre fournisseurs le plus répandu. Les contributions en amont (nouveaux tests, signalements de défauts comme celui de RGW) font partie du travail.

4.7 Les cinq niveaux d'interopérabilité

Le cas d'usage se traite surtout aux niveaux technique et syntaxique. Les cinq niveaux de la note de vision doivent tous être couverts :

Niveau	Exigences pour la sortie d'un stockage objet
Technique	Profil d'interface, limitation de débit, débit sortant, export assisté
Syntaxique	Format du manifeste, encodage des clés, format des métadonnées
Sémantique	Signification des métadonnées (dates d'origine et dates de copie, statut des versions, sémantique de rétention) ; correspondance des classes de stockage
Organisationnel	Procédure de sortie documentée, assistance pendant la transition, interlocuteur, délais (préavis, transition, récupération)
Juridique	Clauses du Data Act, effacement à l'issue de la période de récupération, attestation d'effacement, données soumises à rétention légale, réversibilité des clés de chiffrement

5. État du consensus

Les éléments suivants ont fait consensus lors de l'atelier du 9 octobre 2026, sous réserve de la relecture du groupe de travail :

Sujet	État
Profil	Sous-ensemble documenté de l'API S3, versionné ; niveau Core en première version, autres niveaux en périmètres avancés ; déclaration des fonctions prises en charge avec preuve
Manifeste	JSON Lines décrit par un JSON Schema ; empreinte SHA-256 de l'objet complet ; valeurs d'origine préservées sans fonctionnement natif
Intégrité	L'ETag n'est pas une preuve ; le CTK calcule une empreinte indépendante
Droits	Reconstruction des droits convenus, comparés par leur effet ; critère obligatoire
Scorecard	Cinq critères obligatoires sans compensation, en première étape ; mesures publiées à part ; pas de score global pondéré ; pas de seuil universel de débit
Scénarios	CTK-01 à CTK-05 adoptés ; CTK-06 à affiner ; quatre verdicts
Outillage	Outil de transfert indépendant ; deux implémentations de référence indépendantes
Opposabilité	Le CTK fournit une preuve technique ; l'usage contractuel demande une rédaction propre

L'atelier a classé à part, comme périmètres avancés, les versions historiques et les marqueurs de suppression, le stockage archivé, Object Lock, les règles de cycle de vie, la migration pendant les écritures, les dépendances aux services de gestion de clés et la cohérence forte. Il a aussi ouvert deux rubriques à définir : « infrastructure » (localisation, réseau, distribution, disponibilité) et « niveau de conformité ».

6. Feuille de route

L'atelier a proposé le calendrier suivant, qui reste à confirmer par le groupe de travail :

Date proposée	Livrable	État au 11 novembre 2026
Mi-octobre 2026	Compte rendu, contrats V0, backlog	Contrats, scorecard et scénarios rédigés en version de travail, à relire
Fin octobre 2026	Profil technique et jeu de données stabilisés	Jeu de données et schéma du manifeste mis en œuvre dans le CTK S4 ; rédaction du profil Core en cours
Début novembre 2026	TechSprint et CTK minimal	En cours ; CTK-01 à CTK-05 implémentés, essais contre sept implémentations libres
Fin novembre 2026	Essais croisés et rapport d'écarts	À faire : exécutions entre fournisseurs au volume fonctionnel, CTK-05 au volume de charge
Décembre 2026	Profil V0.1 et proposition d'annexe contractuelle	À faire

Au-delà de décembre 2026, le backlog comprend, par priorité décroissante :

1. le profil S4 au niveau Core (opérations, comportements, erreurs, limitation) ;
2. le manifeste d'export (schéma à valider, empaquetage à définir) ;
3. la méthode de mesure et CTK-06 ;
4. la déclaration de conformité du fournisseur, sur le modèle ICS de l'ETSI ;
5. l'extension de CTK-04 aux erreurs serveur et aux identifiants expirés ;
6. les niveaux Versioned et Governance ;
7. les règles de correspondance (classes de stockage, modèle de droits, chiffrement) ;

8. les classes froides et l'export assisté ;
9. l'extension de la méthode à d'autres services.

À plus long terme, les étapes suivantes restent indicatives : profil v1.0 et premières fiches de profil de fournisseurs, puis clauses types pour la commande publique, soumission à l'autorité technique et présentation à l'ARCEP et à la DG CNECT au titre de l'article 35 du Data Act. Cet horizon de 9 à 12 mois correspond à celui du chantier CL4.

7. Points ouverts

Contenu des exigences.

1. Liste des champs de métadonnées obligatoires, transformations admises et place des étiquettes.
2. Format de la matrice d'accès, droits minimaux pour l'export et l'import, correspondances entre modèles IAM. Le CTK S4 propose un format de matrice, à valider.
3. Incidents, délais et erreurs couverts par la reprise ; seuils de reprise, à calibrer après les premiers essais.
4. Contenu de CTK-06, méthode de comparaison des performances et tolérances entre valeurs déclarées et mesurées.
5. Volume du profil de charge, selon les moyens disponibles.
6. Contenu des rubriques « infrastructure » et « niveau de conformité ».
7. Définition opérationnelle de l'équivalence fonctionnelle de l'article 30 du Data Act pour le stockage objet, par exemple « même niveau de profil, mêmes critères obligatoires satisfaits ».

Gouvernance et usage.

8. Gouvernance du profil, processus de révision face aux évolutions de l'API décidées par AWS, liaison avec SNIA.
9. Rédaction de l'annexe contractuelle : profil et version applicables, fonctions obligatoires, méthode de réception, preuves, conditions d'un nouveau test, responsabilités de la source, de la cible et du client, traitement des écarts.
10. Politique de publication des résultats (anonymisés au début ou non) et articulation des niveaux de preuve avec les schémas de qualification existants, dont SecNumCloud.

Moyens et périmètre.

11. Participation des fournisseurs : comptes de test et quotas suffisants pour les mesures de débit. Le CTK n'a de valeur que s'il est exécuté contre des services réels.
12. Coût des essais à grande échelle : stockage, trafic sortant et temps. La méthode doit permettre d'extrapoler à partir de volumes modestes et réserver les essais à pleine échelle à des campagnes ponctuelles.
13. Réversibilité des clés gérées par le fournisseur (KMS), condition de la réversibilité des données chiffrées ; hors de la première version, à inscrire au backlog.
14. Seconde implémentation de référence : correction du défaut de RGW `dbstore` ou cible RGW sur grappe RADOS.

Références

Réglementation

- Règlement (UE) 2023/2854 du 13 décembre 2023 (Data Act), chapitre VI. https://eur-lex.europa.eu/eli/reg/2023/2854/oj
- Règlement (UE) 2018/1807 sur la libre circulation des données à caractère non personnel. https://eur-lex.europa.eu/eli/reg/2018/1807/oj
- Loi n° 2024-449 du 21 mai 2024 visant à sécuriser et à réguler l'espace numérique (SREN).
- Circulaire du Premier ministre du 5 février 2026 sur la commande publique numérique.
- ANSSI, référentiel SecNumCloud.

Normes et spécifications

- ISO/IEC 19941:2017, *Cloud computing — Interoperability and portability*.
- ISO/IEC 19086-1 à 4, *Cloud computing — Service level agreement (SLA) framework*.
- ISO/IEC 17826, *Cloud Data Management Interface (CDMI)* (SNIA).
- RFC 8493, *The BagIt File Packaging Format (V1.0)*. https://www.rfc-editor.org/rfc/rfc8493
- RFC 9700, *Best Current Practice for OAuth 2.0 Security*. https://www.rfc-editor.org/rfc/rfc9700
- SNIA, *When AWS S3 Keeps Changing, Who Keeps Up?*, juin 2026. https://www.snia.org/blog/2026/when-aws-s3-keeps-changing-who-keeps
- SNIA, *Report from September "25 Cloud Object Storage Plugfest*. https://www.snia.org/sniadeveloper/session/19296
- OpenStack Swift API. https://docs.openstack.org/swift/latest/
- Kubernetes COSI (*Container Object Storage Interface*). https://github.com/kubernetes-sigs/container-object-storage-interface
- A. Brandes, *The Substitution Escrow Threshold: When « Compatible With » Becomes Safe Enough to Buy*, arXiv:2608.21221, août 2026. https://arxiv.org/abs/2608.21221
- SWIPO, codes de conduite IaaS et SaaS. https://swipo.eu/
- Amazon S3 API Reference (interface de fait). https://docs.aws.amazon.com/AmazonS3/latest/API/

Outils

- CTK S4. https://git.sr.ht/~sfermigier/s4-ctk
- ceph/s3-tests. https://github.com/ceph/s3-tests
- rclone. https://rclone.org/
- Garage (Deuxfleurs). https://garagehq.deuxfleurs.fr/
- Ceph RGW. https://docs.ceph.com/en/latest/radosgw/
- MinIO mint et warp. https://github.com/minio/mint, https://github.com/minio/warp

Autres notes

- Synthèse du projet S4
- Contrats comportementaux
- Scorecard de réversibilité
- Scénarios de test
- Implémentation du CTK S4
- Participer aux essais croisés
- Par où commencer
- Proposition de doctrine pour Open Interop
- Note de vision : interopérabilité opposable, cinq niveaux
- État de l'art

- Fiche transverse
- Fiche cloud : chantier CL4, plan d'action CSF Cloud
- Note sur le RGI