

Collecte de données environnementales par l'Arcep - Nouveaux indicateurs pour l'année 2026

Consultation publique – Contribution de France Datacenter

France Datacenter soutient l'objectif visant à établir un bilan plus fin de l'empreinte environnementale du numérique. La filière est consciente du besoin de transparence sur les consommations des datacenters à la fois pour les acteurs institutionnels, via l'Arcep, et pour le public. Les indicateurs collectés durant les années précédentes fournissent déjà un cadre intéressant et suffisant, et nous saluons la décision de l'Arcep de ne pas ajouter de nouveaux indicateurs pour les centres de données. En effet, il est important de rappeler que :

- La collecte de données environnementales de l'Arcep doit s'inscrire dans une forme de stabilité pour permettre d'analyser les évolutions des indicateurs sur le long terme ;
- La collecte doit s'inscrire dans la même logique et la même temporalité que la collecte européenne dans le cadre de la directive sur l'efficacité énergétique. Nous saluons l'avancée réalisée sur ce point avec les différents interlocuteurs français.

Nous souhaitons cependant alerter l'Arcep sur plusieurs points d'attention pour nous en lien avec cette nouvelle collecte ciblant les fournisseurs de cloud :

- **Délais liés à la collecte**
 - La publication de la décision de collecte de l'Arcep a lieu généralement tardivement (en fin d'année), pour une remontée de données fin mars. Ces délais sont trop courts pour nos membres. Un délai de trois mois ne permet pas d'initier les processus internes nécessaires à la mise en place d'un nouveau système de collecte, sans parler du temps nécessaire pour former les équipes, mettre en place les outils de mesure et de collecte appropriés et valider la fiabilité des données. Cette approche risque de générer des coûts supplémentaires dommageables pour les opérateurs, mais également de produire des données de moins bonne qualité qui ne serviront pas l'objectif de transparence environnementale visé ;
 - Nous appelons donc l'ARCEP à adopter des pratiques de consultation plus adaptées aux contraintes opérationnelles du secteur, en prévoyant des délais raisonnables entre la consultation, la publication de la décision et le début effectif de la collecte. Un minimum de **six mois** après la publication de la décision finale est nécessaire pour permettre une mise en œuvre efficace de nouvelles obligations déclaratives.
- **Proportionnalité de la collecte**
 - Chaque année, la collecte de l'Arcep est enrichie pour ajouter de nouvelles informations. Si certaines données sont importantes à collecter pour apprécier et mesurer l'empreinte environnementale des centres de données et du secteur de manière générale, il convient de limiter la collecte aux données les plus importantes et que chaque nouvelle information demandée soit largement justifiée. En effet, la collecte de données constitue une charge supplémentaire importante pour les opérateurs, surtout si elles sont très détaillées.
- **Périmètre de la collecte**
 - Une grande partie des données sollicitées est impossible à fournir. A titre d'exemple, les fournisseurs de cloud ne disposent pas d'une traçabilité de la finalité associée à l'achat de certains services. Les données demandées sont donc matériellement

impossibles à récupérer : un fournisseur ne sait pas si un serveur équipé de GPUs est utilisé pour du gaming, de l'IA générative ou de l'inférence.

- **Risques liés au double comptage**

- L'ajout d'un questionnaire sur les fournisseurs de cloud avec de très nombreux indicateurs ajoute une forte charge administrative pour les acteurs qui opèrent leurs datacenters et fournissent également des services cloud, puisqu'ils sont soumis à une double déclaration ;
- En lien avec le point précédent, l'extension du périmètre aux services d'informatique en nuage soulève des questions importantes quant à la délimitation des responsabilités et la potentielle duplication des efforts de reporting. En effet, il est nécessaire d'opérer une clarification quant à la responsabilité des fournisseurs de cloud et des opérateurs de centres de données liée au risque de double comptage. Les opérateurs de centres de données sont les mieux placés pour mesurer et rapporter les métriques environnementales liées aux infrastructures qu'ils exploitent - notamment la consommation électrique globale, l'utilisation d'eau pour le refroidissement et les émissions de gaz à effet de serre associées. Demander aux clients hébergés de reporter ces mêmes indicateurs créerait non seulement une redondance administrative contre-productive, mais risquerait également de produire des données moins précises basées sur des estimations approximatives plutôt que sur des mesures directes. Nous recommandons vivement que les indicateurs environnementaux liés aux infrastructures restent sous la responsabilité exclusive des opérateurs de centres de données.

- **Enjeux de confidentialité et de sécurité**

- La divulgation des localisations précises des centres de données, l'identification des relations contractuelles entre opérateurs et clients et du trafic des données entrant et sortant des équipements informatiques soulève de sérieuses préoccupations en matière de sécurité et de confidentialité commerciale. Ces informations sont particulièrement sensibles dans notre secteur où la discrétion est essentielle tant pour des raisons de sûreté que de concurrence. La publication de ces données, même de manière agrégée, doit être soigneusement encadrée pour éviter toute possibilité d'identification indirecte des acteurs ou des relations commerciales. La plupart des données collectées dans le cadre de cette enquête sont en réalité protégées par le secret des affaires (voir ci-après).

- **Alignement avec le cadre européen**

- Le calendrier proposé pour janvier 2026 nous semble précipité au regard des développements réglementaires européens en cours. Le nouveau système européen de notation des centres de données, qui sera finalisé au premier trimestre 2026, établira des définitions et méthodologies harmonisées. Il est indispensable d'attendre ces standards pour garantir la cohérence des données collectées entre l'Union européenne et la France. La multiplication de cadres de reporting non alignés augmenterait inutilement la charge administrative des opérateurs sans apporter de bénéfice tangible en termes de transparence environnementale.

Recommandations pour une mise en œuvre efficace et garantir une collecte de données pertinente :

1. Un report de la mise en œuvre à janvier 2027 pour permettre l'alignement avec les standards européens ;
2. Une clarification explicite des responsabilités entre opérateurs d'infrastructures et clients hébergés, ainsi que le maintien et l'exclusivité du reporting des métriques environnementales des infrastructures au niveau des opérateurs de centres de données ;

3. L'ajustement du périmètre des données demandées, certaines étant impossibles ou trop complexes à recueillir ;
4. Une notice expliquant les raisons pour lesquelles l'Arcep souhaite collecter ces nouvelles données, car il nous semble qu'elles ne participent pas à améliorer la transparence des données environnementales ;
5. L'établissement de garanties solides concernant la confidentialité des données sensibles.

2.7 Objet de l'annexe F – Enquête relative aux impacts environnementaux des fournisseurs de services d'informatique en nuage

2.7.1 Personnes soumises à l'annexe F

Sont soumis à la collecte de données de l'annexe F de la présente décision les fournisseurs de services d'informatique en nuage qui sont propriétaires ou exploitants d'équipements informatiques, quelle que soit la nature du service d'informatique en nuage qu'ils fournissent, qu'ils soient propriétaires ou non de leurs propres centres de données. Par mesure de proportionnalité, sont soumis à l'annexe F les fournisseurs de services d'informatique en nuage dont le chiffre d'affaires en France est égal ou supérieur à 10 millions d'euros hors taxes ou dont la demande de puissance des technologies de l'information installées et exploitées est supérieure ou égale à 100 kW. Pour rappel, les fournisseurs de services d'informatique en nuage propriétaires de leurs propres centres de données sont également soumis à la collecte de données de l'annexe B de la présente décision dans les conditions prévues par cette dernière.

Question 9. Quelles remarques souhaitez-vous formuler sur le seuil à compter duquel l'obligation de communication d'informations s'applique aux fournisseurs de services d'informatique en nuage ?

Le seuil proposé est identique à celui appliqué aux centres de données, par souci de cohérence. Par ailleurs, il faut clarifier la répartition des responsabilités entre opérateurs de colocation et fournisseurs de cloud. Il existe un risque de double comptage entre les données déjà collectées pour les centres de données et celles à fournir par les fournisseurs de cloud. Nous suggérons donc de retirer les indicateurs liés aux sites (qui relèvent du volet "centres de données") et de faire communiquer directement par les fournisseurs de cloud les informations sur leurs équipements en colocation, plutôt que par les opérateurs.

Question 10. Quelles remarques souhaitez-vous formuler sur la représentativité du marché considéré, compte tenu du seuil proposé ?

Le seuil nous convient, il couvre la plupart des cloud providers.

2.7.2 Nature des données collectées à l'annexe F

Les données requises portent sur :

- Les émissions de gaz à effet de serre en France, associées à l'activité de fournisseur de services d'informatique en nuage : [ces données ne doivent pas faire l'objet d'un double comptage et devraient être collectées par les centres de données](#) ;
- Pour chaque centre de données hébergeant les équipements informatiques qui supportent les services du fournisseur de services d'informatique en nuage :
 - o Le nom du centre, de son opérateur et de la commune dans laquelle il se situe ;

- La demande de puissance des équipements informatiques installées pour leurs services d'informatique en nuage : ces données ne doivent pas faire l'objet d'un double comptage et devraient être collectées par les opérateurs de centres de données ;
- Les consommations d'électricité et d'eau associées aux équipements informatiques utilisés pour leurs services d'informatique en nuage : ces données ne doivent pas faire l'objet d'un double comptage et devraient être collectées par les opérateurs de centres de données ;
- Le trafic de données entrant et sortant des équipements informatiques utilisés pour leurs services d'informatique en nuage : comme indiqué à de multiples reprises, le trafic de données relève des fournisseurs d'accès à internet (FAI) et du choix des clients finaux, et présente peu d'intérêt en matière environnementale car il ne reflète pas l'intensité de calcul réelle qui détermine la consommation énergétique. Le volume de données sortant est déconnecté de la valeur d'usage ;
- Le nombre de serveurs en service au 31/12 de l'année considérée et le nombre de serveurs mis en service dans l'année. Les serveurs sont segmentés selon qu'ils disposent ou non d'un processeur graphique (GPU) - cette métrique n'est pas pérenne car les serveurs modernes peuvent être partitionnés en plusieurs machines virtuelles, le comptage ne reflète pas l'efficacité réelle ; et l'évolution rapide des technologies rend ces catégories rapidement obsolètes :
 - Les serveurs avec GPU sont catégorisés en fonction de l'offre la plus exigeante en puissance de calcul que l'un de leurs GPU est capable de supporter (GPU pour l'entraînement de modèles d'intelligence artificielle (IA) générative, GPU d'inférence d'IA générative, autres types de GPU et au moins un processeur (CPU) optimisé pour l'IA générative, autres types de GPU et pas de CPU optimisé pour l'IA générative) ;
 - Les serveurs sans GPU sont catégorisés selon qu'ils disposent ou non d'au moins un CPU optimisé pour l'IA générative ;
- Le nombre total de GPU et CPU en service au 31/12, segmentés en fonction de leur usage possible (GPU : entraînement de modèles d'IA générative, inférence d'IA générative, autres usages et CPU : optimisés pour l'IA générative ou non) : la récolte de ces données n'a pas d'objectif clair et la méthodologie n'est pas pérenne. En effet, ils reposent sur des définitions instables et non standardisées, ce qui entraînerait des rapports incohérents et non comparables entre les fournisseurs. Le matériel peut être réaffecté à différents usages au fil du temps, et la distinction entre équipements IA et non-IA est floue. De plus, cette approche ignore l'existence d'autres technologies comme les circuits intégrés pour applications spécifiques (ASIC) et ne tient pas compte de la virtualisation qui permet à un serveur physique d'héberger plusieurs machines virtuelles.
- Le nombre de serveurs mis hors service dans l'année segmentés selon : ces données sont confidentielles et pourraient être mal interprétées sans contexte sur les besoins d'évolution technologique et de modernisation :
 - Leur destination (revente/recyclage) : le pourcentage de serveurs mis hors service soumis au remarketing logistique local (RLL) pour évaluation en vue de revente ou de recyclage peut être communiqué, mais pas les chiffres absolus ;
 - Le type de GPU le plus performant présent, ou, s'ils n'en ont pas, selon la présence d'un CPU optimisé pour l'IA générative : les catégories manquent de définitions standardisées et le matériel peut être réaffecté à différents usages, rendant cette classification peu pertinente ;
- La durée moyenne d'utilisation des serveurs mis hors service et recyclés dans l'année : ces données sont très confidentielles et peu fiable car les équipements sont

fréquemment réaffectés ou mis à niveau avant leur fin de vie, rendant difficile le calcul d'une durée d'utilisation moyenne significative ;

- Le nombre d'équipements de stockage en service au 31/12 et le nombre d'équipements de stockage mis en service dans l'année, segmentés selon le dispositif de stockage : HDD ou SSD : ces données sont confidentielles et ne reflètent pas l'efficacité réelle ni les progrès technologiques en matière de densité de stockage ;
- Le nombre d'équipements de stockage mis hors service dans l'année segmentés selon :
 - Leur destination (revente/recyclage) : le pourcentage de serveurs mis hors service soumis à RLL pour évaluation en vue de revente ou de recyclage peut être communiqué, mais pas les chiffres absolus ;
 - Le dispositif de stockage (HDD/SSD) ;
- La durée moyenne d'utilisation des équipements de stockage mis hors service et recyclés dans l'année : ces données sont confidentielles et la métrique est peu fiable car les équipements peuvent être réaffectés plusieurs fois avant leur fin de vie effective.

Question 11. Quelles remarques souhaitez-vous formuler sur les données qu'il est envisagé de collecter ?

Ces données ne participent pas à améliorer la transparence de l'information sur l'empreinte environnementale du numérique, qui est l'objet de cette collecte. Par ailleurs, aucune justification ne vient étayer cette demande très détaillée de données sensibles. Enfin, elles sont couvertes par le secret des affaires (article L 151-1 du code du commerce) et leur divulgation pourrait compromettre la position concurrentielle des entreprises. La composition détaillée des flottes d'équipements, par exemple, peut révéler des stratégies d'investissement et d'innovation. Les volumes de trafic peuvent indiquer l'ampleur des activités et la base client d'une entreprise. La divulgation des détails au niveau des sites peut exposer les relations commerciales et les capacités opérationnelles, d'où l'importance de capturer un maximum d'acteurs pour que la consolidation ne représente pas de risque d'authentification des données.

Question 12. Quelles difficultés d'ordre méthodologique / opérationnel ou limites ces indicateurs pourraient-ils, selon vous, entraîner ?

Comme indiqué en avant-propos, l'ajout d'un questionnaire sur les fournisseurs de cloud avec de très nombreux indicateurs ajoute une forte charge administrative pour les acteurs. De plus, il convient de souligner que les fournisseurs de cloud, de manière générale et dans leur système d'information, ne peuvent pas retracer la finalité associée à l'achat de certains services cloud. Les données demandées sont aujourd'hui matériellement impossibles à récupérer, puisque le fournisseur ne sait pas si un serveur avec des GPUs est utilisé pour du gaming, de l'IA générative ou de l'inférence.

Par ailleurs, l'attribution de certaines mesures au niveau des sites (comme la consommation d'eau) à des équipements spécifiques s'avère complexe et nécessiterait des méthodes d'estimation précises et des directives bien définies.

Enfin, le partage de données sans contexte adéquat soulève des préoccupations sérieuses en matière de sécurité opérationnelle. En effet, des informations brutes sur la capacité des infrastructures, sans explications sur les exigences de redondance, les prévisions de croissance ou les dispositifs de tolérance aux pannes, pourraient être mal interprétées et potentiellement exploitées de manière

préjudiciable. Cette divulgation pourrait exposer les vulnérabilités des systèmes et compromettre la résilience des infrastructures numériques critiques.

2.7.3 Utilisation des données collectées à l'annexe F

Les données relatives à l'annexe F de la présente décision seront principalement utilisées pour l'élaboration et la publication de statistiques. En particulier, elles seront utilisées pour rendre compte de l'empreinte environnementale des fournisseurs de services d'informatique en nuage dans le rapport d'activité de l'Arcep tel qu'indiqué dans l'article L. 135 du CPCE.

Question 13. Quelles remarques souhaitez-vous formuler concernant les modalités de restitution des indicateurs par l'Arcep ?

La publication de ces données, même de manière agrégée, doit être soigneusement encadrée pour éviter toute possibilité d'identification indirecte des acteurs ou des relations commerciales. La restitution devra être accompagnée d'un contexte présentant la complexité de l'infrastructure pour éviter les interprétations erronées. La plupart des données collectées dans le cadre de cette collecte sont en réalité protégées par le secret des affaires.

Question 14. Quelles autres remarques souhaitez-vous formuler sur l'ensemble des indicateurs présents dans la décision de collecte de données environnementales ?

Au-delà de ces nouveaux indicateurs, nous souhaitons signaler à l'Arcep un problème de terminologie relatif à la collecte de données sur l'eau. En effet, le rapport de l'Arcep mentionne deux indicateurs : les prélèvements et les consommations. Il s'avère que ces deux notions sont définies précisément en hydrologie :

- Les prélèvements désignent les "volumes d'eau douce extraits définitivement ou temporairement d'une source souterraine ou de surface et transportés sur leur lieu d'usage." Après utilisation, une fois assainie, l'eau peut être à nouveau prélevée (30 milliards de m3 en France en 2020) ;
- Les consommations désignent quant à elles « la part des prélèvements ne retournant pas au milieu naturel » (4,4 milliards de m3 en France en 2020).

Or, l'Arcep parle de « consommation » dans une acception qui n'est pas la bonne puisqu'elle désigne dans son rapport l'eau utilisée par les centrales nucléaires pour alimenter les centres de données. Il ne s'agit donc en aucun cas d'une consommation au sens évoqué plus haut, tout au plus une « consommation indirecte » qui pourrait s'apparenter à un « scope 3 » de l'eau. Cette notion pourrait être intéressante mais n'est, à notre connaissance, utilisée dans aucun autre secteur d'activité et mériterait d'être définie et objectivée. Cette notion gonfle artificiellement les données de consommation sur l'eau des datacenters, qui sont en réalité extrêmement faibles puisque l'eau prélevée n'est qu'en partie consommée.

Au nom de la clarté et de l'intelligibilité des données transmises dans le rapport de l'Arcep, nous appelons donc à redéfinir correctement les termes utilisés.

À propos de France Datacenter

France Datacenter est la seule association des acteurs de l'écosystème des datacenters en France. Elle représente et assure la promotion de la filière comme socle de la performance et de la fiabilité de l'économie numérique. L'association est aujourd'hui constituée d'une centaine de sociétés qui conçoivent, construisent et exploitent les datacenters. Acteur naturel de la réflexion sur les enjeux des industries numériques, France Datacenter favorise la formalisation des savoirs, le développement des compétences et l'adoption des meilleures pratiques par les professionnels, afin d'améliorer la compétitivité de la filière française face à ses concurrents européens