



**VERTIV™**

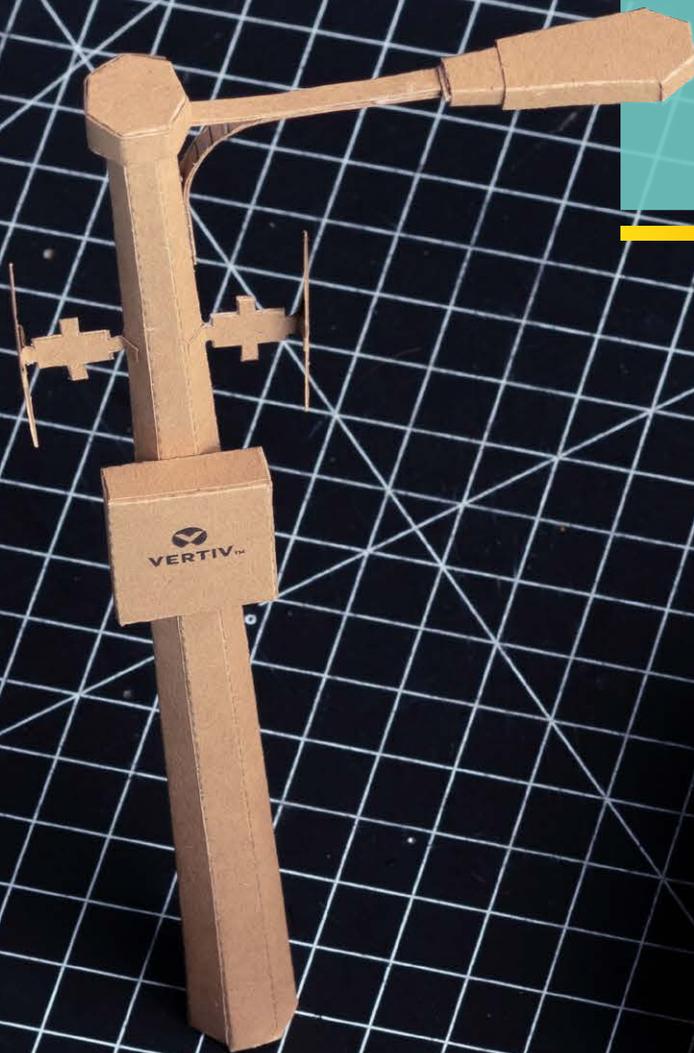
**RAPPORT SPÉCIAL**

**Archétypes Edge 2.0**

**Modèles d'infrastructure**

**Edge prêts à être déployés**

Concevoir l'architecture de l'infrastructure edge computing optimale pour votre entreprise



Développé avec STL Partners

**STL PARTNERS**



## Résumé

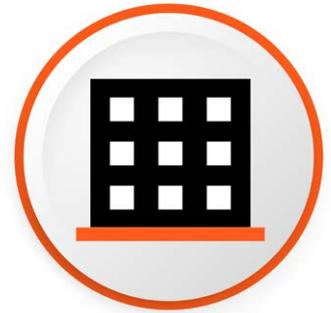
L'infrastructure physique est essentielle dans toute stratégie Edge Computing. L'équipement électrique, de refroidissement et de confinement, ainsi que l'informatique qu'il prend en charge, fournit la base sur laquelle les applications peuvent fonctionner et permet d'innombrables cas d'utilisation Edge.

Faire le bon choix en matière d'infrastructure physique est encore plus important au niveau de l'Edge, étant donné que de nombreux déploiements se situent dans des endroits où une assistance et une protection supplémentaires sont requises. S'y retrouver en termes d'infrastructure Edge est également plus compliqué en raison des définitions larges et variées de ce qu'est l'Edge. Ces facteurs compliquent la tâche des 49%<sup>1</sup> d'entreprises explorant les déploiements Edge Computing. Elles doivent prendre des décisions concernant la meilleure façon d'utiliser l'infrastructure existante et savoir où investir aujourd'hui afin de pouvoir répondre aux besoins de demain. Heureusement, il existe un écosystème de fournisseurs, d'intégrateurs système et d'autres partenaires channel ayant l'expérience et l'expertise requises dans les déploiements Edge pour fournir un soutien.

S'appuyant sur les travaux de Vertiv sur les archétypes Edge<sup>2</sup> qui ont fourni une classification des cas d'utilisation Edge, ce rapport va plus loin en définissant quatre modèles d'infrastructure Edge distincts. Le cadre a été développé sur la base d'entretiens avec une multitude de professionnels sectoriels, d'experts en datacenters, de fournisseurs de solutions et d'organismes industriels dans des applications telles que ville intelligente, santé, fabrication et retail. Grâce à une analyse approfondie des besoins en Edge Computing de différents secteurs et cas d'utilisation, les modèles d'infrastructure Edge Computing suivants ont été définis :

<sup>1</sup> Enquête STL Partners auprès de 699 professionnels sectoriels dans le monde entier, allant des secteurs de la fabrication, du retail, de la santé, du transport et de la logistique, mai 2021

<sup>2</sup> Définition des Quatre Archétypes Edge et leurs Exigences Technologiques



Device Edge	Micro Edge	Datacenter Edge distribué	Datacenter Edge régional
<ul style="list-style-type: none"> <li>• Sur dispositif</li> <li>• Attaché ou intégré</li> <li>• À l'extérieur (p. ex., lampe de rue) ou à l'intérieur (p. ex., équipement de fabrication)</li> </ul>	<ul style="list-style-type: none"> <li>• Petit nombre de serveurs ou racks</li> <li>• 0–4 racks</li> <li>• Sur le site de l'entreprise (p. ex., retail, usine, armoire IT, municipalités)</li> </ul>	<ul style="list-style-type: none"> <li>• Petit datacenter</li> <li>• 5–20 racks</li> <li>• Sur le site de l'entreprise (p. ex., entrepôt, bureau), site de télécommunications, parking</li> </ul>	<ul style="list-style-type: none"> <li>• Datacenter de taille moyenne</li> <li>• Plus de 20 racks</li> <li>• Site régional (p. ex., ville de niveau 2 ou 3)</li> </ul>

## Données clés

- L'infrastructure Edge Computing n'agira pas comme substitut au cloud. Il est estimé que le nombre total de sites Edge connaîtra une croissance de 226 %<sup>3</sup> entre 2019 et 2025. De même, le cloud continuera de croître à un taux de croissance annuel moyen de 10 %<sup>4</sup>.
- Les États-Unis ouvrent la voie avec des initiatives Edge et on estime qu'ils sont le plus grand marché Edge Computing<sup>5</sup>, sous l'impulsion de secteurs clés tels que la fabrication.
- Les déploiements Edge Computing les plus développés sont ceux alignés sur l'archétype Edge Sensibilité à la latence humaine (p. ex., cloud gaming), suivis par Grand volume de données (p. ex., analyse vidéo) et Sensibilité à la latence machine à machine (p. ex., transactions boursières). Les cas d'utilisation des archétypes critiques pour la vie humaine (p. ex., voitures autonomes) sont encore principalement à une étape d'exploration ou de validation de principe.
- La plupart des cas d'utilisation d'archétype critiques pour la vie humaine utiliseront le modèle d'infrastructure Device Edge à moyen terme, tandis que les cas d'utilisation Grand volume de données, Sensibilité à la latence humaine, Sensibilité à la latence machine à machine accéléreront la transition des modèles d'infrastructure de type Datacenter Edge régional vers ceux de types Micro Edge et Datacenter Edge distribué à court terme.
- La coordination des nombreux composants du Edge Computing (logiciel, matériel, infrastructure, etc.) est difficile et requiert un écosystème de partenaires pour soutenir les 66% des entreprises qui préfèrent disposer d'une solution Edge complète provenant d'un fournisseur principal unique.

<sup>3</sup> Datacenter 2025 : Plus près du Edge

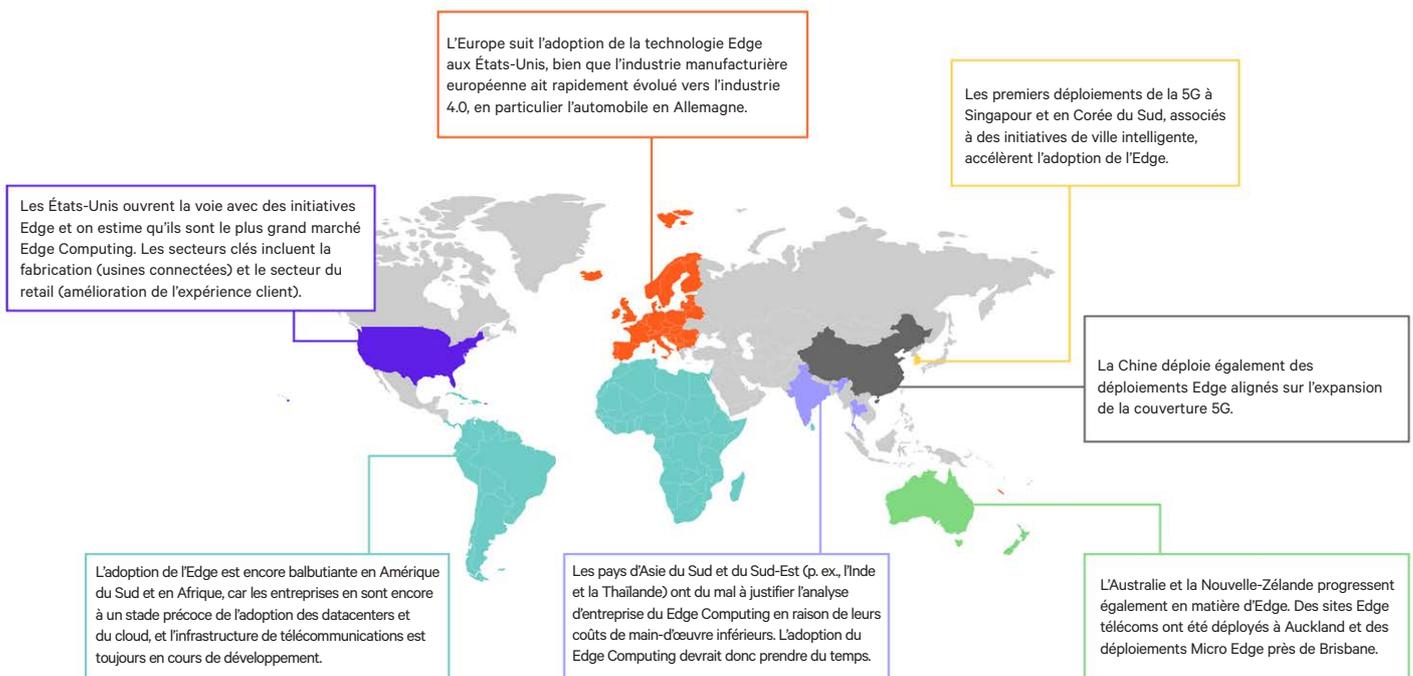
<sup>4</sup> Technavio, 2021

<sup>5</sup> Marché du Edge Computing – Prévisions mondiales d'ici 2025

## Introduction : L'état de l'infrastructure Edge aujourd'hui

Il y a vingt ans, le marché des datacenters s'est tourné vers l'informatique centralisée pour améliorer l'efficacité du traitement des données. Maintenant, le pendule oscille dans l'autre sens vers l'Edge Computing. Edge Computing fait référence aux capacités de calcul et de stockage se situant entre des datacenters centralisés et les utilisateurs finaux, l'appareil ou la source de données. D'une part, l'Edge Computing peut être considéré comme une alternative au cloud et aux datacenters centraux, lorsque ces options ne sont pas en mesure de répondre aux exigences de latence, ou lorsqu'il est trop coûteux de transférer des volumes élevés de données sur de longues distances. D'un autre côté, l'Edge Computing est également un moteur de l'adoption du cloud. Un site Edge peut agir comme une plateforme intermédiaire pour les données qui sont finalement envoyées au cloud aux fins de traitement, stockage ou d'analyse sur le long terme.

Au cours des deux dernières années, l'adoption du Edge Computing a considérablement augmenté, parallèlement à la croissance continue du cloud. Selon une étude récente de STL Partners, 49 % des entreprises de secteurs spécifiques explorent activement l'Edge Computing<sup>6</sup> et on estime que le nombre total de sites Edge augmentera de 226 % entre 2019 et 2025<sup>7</sup>. Pour autant, l'adoption varie selon les zones géographiques. Cela est dû au niveau de maturité des technologies connexes (p. ex., l'intelligence artificielle), à l'infrastructure de télécommunications existante, à la politique de services publics et à l'importance de certains secteurs dans le pays. Par exemple, l'industrie manufacturière stimule l'adoption du Edge Computing aux États-Unis et en Allemagne, et devrait représenter la plus grande part des dépenses Edge des entreprises européennes en 2021<sup>8</sup>.



Les entreprises considèrent l'Edge Computing comme un facteur clé pour surmonter les défis liés à la sécurité et à la fiabilité des données, en plus d'améliorer les performances des applications. Par exemple, de grands clusters de datacenters pourraient devenir des cibles d'attaques privilégiées. La division du cœur de réseau en plusieurs sites Edge peut coûter plus cher par kW, mais élimine la menace de déni de service simultané. L'Edge Computing promet également de bénéficier à une large plage de secteurs d'activité dans un ensemble diversifié de cas d'utilisation. Du cloud gaming aux smart grids pour les réseaux de distribution d'électricité en passant par les robots autonomes dans les environnements industriels, tous ces cas d'utilisation ont quelque chose à gagner du traitement des données plus près du dispositif terminal. Les adoptants précoces mettent déjà en œuvre des solutions innovantes, passant des validations de principe et des pilotes initiaux à des déploiements multisites à grande échelle. Lloyds Register, une entreprise de services maritimes, a par exemple déployé l'Edge Computing sur des flottes de navires<sup>9</sup> afin d'optimiser la consommation de carburant grâce à des données. L'adoption du Edge Computing sera également soutenue par un écosystème croissant de fournisseurs, d'intégrateurs système et d'autres acteurs en matière de distribution. La nature distribuée du Edge Computing exige un réseau d'acteurs Edge ayant la portée suffisante et la capacité de déployer, de servir et de prendre en charge l'infrastructure Edge.

<sup>6</sup> Enquête STL Partners auprès de 699 professionnels sectoriels dans le monde entier, allant des secteurs de la fabrication, du retail, de la santé, du transport et de la logistique, mai 2021

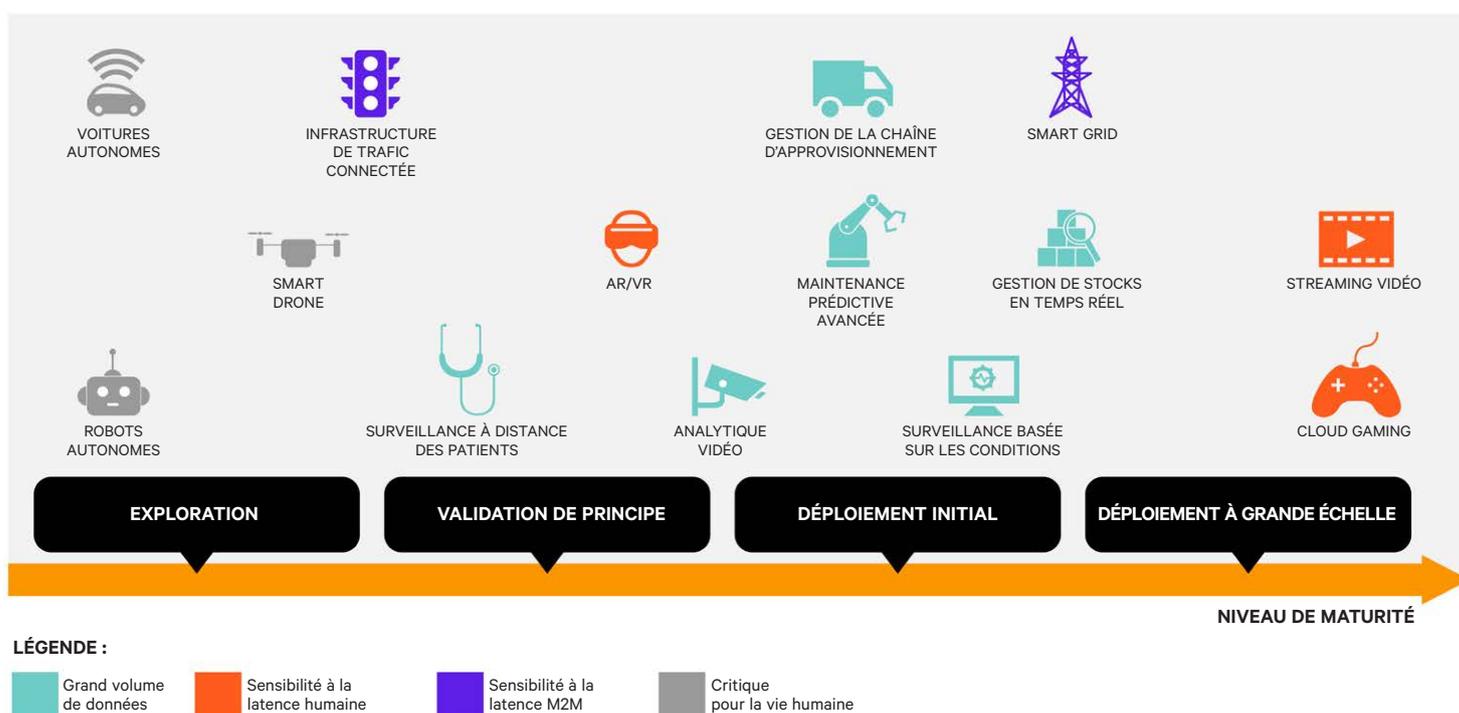
<sup>7</sup> Datacenter 2025 : Plus près du Edge

<sup>8</sup> Worldwide Edge Spending Guide de International Data Corporation (IDC)

En 2018, Vertiv a publié un rapport, Définition des quatre archétypes Edge et de leurs exigences technologiques, qui fournissait un cadre de référence pour la catégorisation des cas d'utilisation. Ces quatre archétypes ont aidé les entreprises et les opérateurs de datacenters Edge à mieux comprendre les exigences sous-jacentes communes à des cas d'utilisation similaires. Les quatre archétypes sont :

- **Grands volumes de données** : Cas d'utilisation pour lesquels la quantité de données complique le transfert via le réseau directement vers le Cloud, ou du Cloud au point d'utilisation, en raison de problèmes de volume de données, de coût ou de bande passante.
- **Sensibilité à la latence humaine** : Cas d'utilisation dans lesquels les services sont optimisés pour une consommation humaine, ou pour améliorer l'expérience des utilisateurs au moyen de services technologiques. La vitesse est la caractéristique principale de ce cas d'utilisation, car les retards dans la livraison des données ont un impact direct sur l'expérience utilisateur.
- **Sensibilité à la latence de machine à machine (M2M)** : Cas d'utilisation pour lesquels les services sont optimisés pour une consommation machine à machine. Comme les machines peuvent traiter les données beaucoup plus rapidement que les humains, la vitesse est la caractéristique principale ici (et les conséquences de l'incapacité à fournir des données dans les délais requis peuvent être plus élevées que pour les cas d'utilisation sensibles à la latence humaine).
- **Critique pour la vie humaine** : Cas d'utilisation ayant un impact direct sur la santé et la sécurité humaines. La vitesse et la fiabilité sont donc primordiales.

Des entretiens avec des experts des différents secteurs d'activité et de l'espace des datacenters ont révélé que les archétypes varient concernant leur niveau de maturité. **Les cas d'utilisation Edge sensibles à la latence humaine** (p. ex., le cloud gaming) sont les plus matures et sont déjà au niveau de déploiements à grande échelle. La croissance de la 5G et l'augmentation des déploiements de fibre optique vont accélérer davantage cette maturité. Inversement, les cas d'utilisation **critiques pour la vie humaine** exigeront beaucoup plus de temps concernant l'adoption du Edge Computing. Cela s'explique par le fait qu'ils ont des exigences strictes en matière de latence et de fiabilité et qu'ils ont souvent besoin de modifications réglementaires à mettre en œuvre à grande échelle. Les drones intelligents sont un exemple. Les gouvernements doivent être convaincus que les drones autonomes n'impliqueront aucune menace pour la vie humaine avant de relaxer les restrictions au niveau de l'espace aérien. De même, l'infrastructure de trafic connectée est encore à un stade précoce. Aux États-Unis uniquement, seuls 7 %<sup>10</sup> des feux de circulation sont intelligents.



**LÉGENDE :**

- Grand volume de données
- Sensibilité à la latence humaine
- Sensibilité à la latence M2M
- Critique pour la vie humaine

<sup>9</sup> WWT, 2020 : Trois études de cas en situation réelle expliquant comment les fabricants peuvent maximiser l'Edge Computing

<sup>10</sup> Programme d'entretiens Vertiv – citation de la personne interrogée (Directeur – Expérience IA, constructeur automobile).

## Évolution des cas d'utilisation vers l'infrastructure

Trois ans après la publication du rapport original sur les archétypes, le marché du Edge Computing est toujours en évolution et les entreprises continuent à développer leurs solutions Edge Computing. Les cas d'utilisation ont progressé du concept vers des applications réelles déployées sur le terrain. Ces applications logicielles ont besoin d'une infrastructure adéquate capable de prendre en charge le traitement de données à bande passante élevée et faible latence au niveau du Edge.

Le terme « infrastructure Edge » fait référence à l'infrastructure de calcul physique (serveurs, alimentation électrique, refroidissement, confinement) qui est délibérément positionnée à quelconque endroit entre le dispositif terminal et les datacenters centraux. Cela inclut également l'hébergement des fonctionnalités de calcul sur site, quelque chose qui n'est évidemment pas nouveau pour de nombreuses entreprises. De fait, certaines réinvestissent dans une infrastructure sur site existante (p. ex., serveurs, armoires réseau ou datacenters) pour optimiser les applications et mettre en œuvre de nouveaux cas d'utilisation. Par exemple, un fabricant multinational de pâtes et de papier<sup>11</sup> permet l'emploi d'applications intensives en données telles que la maintenance prédictive avancée en exploitant les datacenters de ses plus grandes usines.

En se fondant sur une définition stricte, une véritable infrastructure Edge doit utiliser une infrastructure IT standard et être configurée selon les principes du cloud pour héberger des applications et des charges de travail natives du cloud. L'infrastructure sur site héritée qui est monolithique ou basée sur du matériel propriétaire n'est pas considérée comme relevant de l'Edge Computing selon cette définition.

À ce jour, le marché manque de clarté concernant la signification d'une infrastructure Edge.<sup>12</sup> Les entreprises clientes veulent adopter des solutions Edge aujourd'hui en étant assurés que ces solutions répondront aux besoins futurs. De même, les opérateurs de datacenter Edge doivent investir dans l'infrastructure actuelle qui prendra en charge les futures applications. Les deux ont besoin de réponses à des questions clés concernant l'infrastructure Edge Computing :

- À quoi ressemble l'Edge en termes d'infrastructure physique ?
- Quels seront les avantages mesurables du déploiement de l'équipement IT plus près des applications ?
- Qui sera le propriétaire et l'exploitant de l'infrastructure Edge Computing ?
- Comment pouvons-nous la mettre en œuvre efficacement et à grande échelle ?

Dans ce livre blanc, nous allons explorer les facteurs clés qui influencent l'infrastructure Edge, ceci incluant le cas d'utilisation, le secteur et l'environnement externe. Dans le cadre de cette recherche, nous avons conduit 22 entretiens avec une multitude de professionnels sectoriels, ceci incluant des entreprises, des experts en datacenters, des fournisseurs de solutions ainsi que des organismes industriels.

## Créer votre Edge : Quatre modèles d'infrastructure Edge fournissent les bases

Vertiv a développé un cadre innovant pour catégoriser l'infrastructure Edge en modèles spécifiques afin d'aider les organisations à prendre des décisions pratiques concernant le déploiement de l'infrastructure physique et du Edge Computing. Le terme « infrastructure » est utilisé à la place de celui de datacenter, car chaque déploiement Edge ne peut pas être simplement décrit comme relevant d'un datacenter en soi.<sup>13</sup> Les modèles aident à aligner la terminologie utilisée lors des discussions concernant l'Edge Computing. Cela englobe la variété des déploiements Edge observés aujourd'hui, ainsi que l'évolution des déploiements attendus au cours des prochaines années.

Les quatre modèles d'infrastructure Edge sont les suivants :

- **Device Edge** : La capacité informatique se situe au niveau du dispositif terminal. Elle est soit intégrée au dispositif (p. ex., une caméra vidéo intelligente avec des fonctionnalités d'intelligence artificielle) soit un facteur de forme autonome « add-on-edge » qui se fixe directement au dispositif (p. ex., un ordinateur Raspberry Pi attaché à un véhicule guidé automatisé). Lorsque la capacité de calcul est intégrée, le matériel IT est entièrement enfermé dans le dispositif, de sorte qu'il n'a pas besoin d'être conçu pour résister aux environnements difficiles. Par exemple, lorsque cette capacité se situe à l'extérieur d'une caméra, un renforcement est requis, mais en cas d'intégration à la caméra, elle se trouve dans un environnement contrôlé, aussi un renforcement n'est pas nécessaire.
- **Micro Edge** : Une petite solution autonome qui peut aller d'un ou deux serveurs à quatre racks. Ce type d'Edge est souvent déployé sur le site d'une entreprise (p. ex., il peut être installé dans l'atelier d'un fabricant, dans l'usine ou dans un back-office). Il peut également être situé sur un site de télécommunications (p. ex., un rack serveur situé au niveau d'une station de base de télécommunications). Le Micro Edge peut être déployé dans des environnements climatisés et non climatisés. Dans les environnements climatisés (p. ex., armoire IT), le Micro Edge ne nécessite pas de refroidissement et de filtration avancés, car les facteurs externes tels que la température et la qualité de l'air sont stables. Dans des environnements non climatisés (p. ex., dans

<sup>11</sup> Personne interrogée dans le cadre du programme de recherche Vertiv, 2021

<sup>12</sup> Infrastructure Edge Computing fait référence à la plateforme IT Edge ainsi qu'aux installations physiques qui la prennent en charge (p. ex., alimentation électrique, refroidissement, sécurité, confinement).

<sup>13</sup> Un environnement de datacenter typique comprend normalement : connexion fibre optique, ASI, refroidissement, sécurité, câblage, faux plancher.

un atelier d'usine), l'équipement est renforcé et le Micro Edge requiert un refroidissement et une filtration spécialisés afin de tenir compte des facteurs externes les plus agressifs (p. ex., températures élevées et poussière).

- **Datacenter Edge distribué** : Un petit datacenter avec moins de 20 racks qui est situé sur le site de l'entreprise, des sites de réseau télécom ou au niveau d'un site régional (p. ex., dans des usines modernes ou de grandes propriétés commerciales).

- **Datacenter Edge régional** : Un site de datacenter situé à l'extérieur des hubs de datacenter centraux. Étant donné qu'il s'agit généralement d'un site spécialement conçu pour héberger l'infrastructure informatique, il partage de nombreuses caractéristiques avec les datacenters Hyperscale (par exemple, il est climatisé et contrôlé, bénéficie d'une sécurité élevée et offre une grande fiabilité).

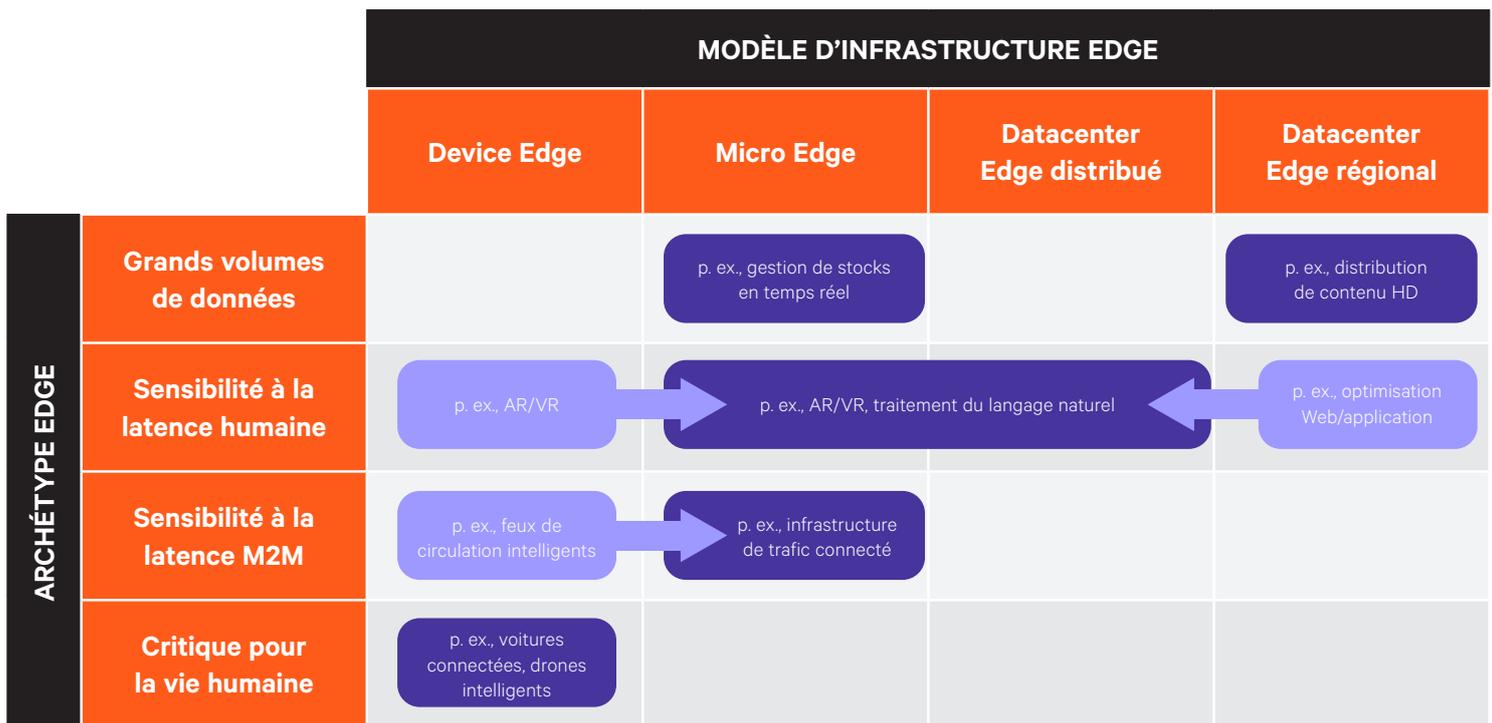


MODÈLE D'INFRASTRUCTURE EDGE				
	Device Edge	Micro Edge	Datacenter Edge distribué	Datacenter Edge régional
<b>Emplacement</b>	Appareils intelligents (p. ex., dans un véhicule, réverbère, IoT)	Site d'entreprise (p. ex., retail, usine, armoire IT, municipalités)	Site d'entreprise (p. ex., entrepôt, bureau), site de télécommunications, parking, ville de niveau 2/3	Ville niveau 2/3 <sup>14</sup>
<b>Nombre de racks</b>	0	0–4 racks	5–20 racks	Plus de 20 racks
<b>Alimentation électrique</b>	Jusqu'à 1 kW	Jusqu'à 20 kW	Jusqu'à 200 kW	Jusqu'à 4 000 kW
<b>Location</b>	Locataire unique	Locataire unique	Locataire unique/Multi-locataires	Multi-locataires
<b>Environnement externe</b>	Contrôlé (à l'intérieur de l'appareil), difficile et conditions extrêmes	Armoire IT, Commercial et bureau, difficile et conditions extrêmes	Difficile et conditions extrêmes, Commercial et bureau, Climatisé et contrôlé	Climatisé et contrôlé
<b>Infrastructure passive</b>	Peut disposer ou non d'alimentation et de filtration, pas de refroidissement, etc.	Dispose d'alimentation avec refroidissement et filtration limités, etc.	Niveau 1+	Niveau 3+
<b>Fournisseur d'infrastructure Edge</b>	Fabricant d'appareils ou solution interne au sein de l'entreprise/du service public	Matériel OEM, fournisseur de datacenter, opérateur télécom ou solution interne au sein de l'entreprise/du service public	Fournisseur de colocation, fournisseur de cloud Hyperscale (cloud public), opérateur télécom	Fournisseur de colocation, fournisseur de cloud Hyperscale (cloud public)
<b>Déploiements attendus</b>	Millions	Centaines de milliers	Milliers	Centaines

\* d'ici 2030 par région majeure

L'identification du modèle d'infrastructure Edge approprié dépend du cas d'utilisation déployé. Étant donné que des cas d'utilisation similaires ont souvent des exigences similaires, il peut être utile de commencer par identifier l'archétype Edge.

Généralement, plus la latence requise est faible, plus l'infrastructure Edge doit être proche de l'appareil terminal. Pour cette raison, les cas d'utilisation **critiques pour la vie humaine** doivent souvent être hébergés sur le **Device Edge**, tandis que les cas d'utilisation avec **grand volume de données** sont souvent hébergés sur site au niveau du **Micro Edge**.



#### LÉGENDE :

- Modèle d'infrastructure qui est généralement déployé aujourd'hui
- Modèle d'infrastructure qui sera le plus déployé à l'avenir selon nous

- Grands volumes de données :** Étant donné que les cas d'utilisation impliquant de grands volumes de données exigent que l'Edge soit proche de la source de données pour éviter des coûts de bande passante élevés, des déploiements sur site sont souhaitables. Un Micro Edge offre un bon équilibre entre une courte distance en matière de transmission de données (limitant ainsi les coûts de bande passante) et de meilleures fonctionnalités de calcul qu'un Device Edge.
- Sensibilité à la latence humaine :** L'archétype sensible à la latence humaine est dominé par les applications grand public (par exemple, optimisation de la vitesse de site Web<sup>15</sup>) pour lesquelles une solution Edge sur site ne constitue pas une option. Pour cette raison, la plupart des cas d'utilisation sensibles à la latence humaine sont aujourd'hui hébergés dans des datacenters Edge régionaux. Cependant, du fait que la latence doit passer à une plage inférieure à 10 millisecondes et que les datacenters Edge deviennent plus disponibles au niveau des sites d'accès<sup>16</sup>, les datacenters Edge distribués constitueront une option favorable. Les applications sensibles à la latence humaine relative aux activités (par exemple, AR/VR) sont généralement hébergées sur un Device Edge aujourd'hui afin de répondre aux exigences en termes de latence, mais passeront au modèle Micro Edge sur site, car ce type d'Edge est de plus en plus déployé par les entreprises.
- Sensibilité à la latence M2M :** Les machines peuvent traiter les données beaucoup plus rapidement que les humains, la vitesse constitue donc l'exigence déterminante des applications sensibles à la latence machine à machine. Le Device Edge répond à ces exigences de latence, pour autant une transition vers le modèle Micro Edge sera opérée, car l'adoption du Edge d'entreprise devient plus répandue, en particulier pour les appareils machine à machine qui sont trop petits ou à faible coût pour justifier un Device Edge. Par exemple, dans le secteur de la fabrication, les fournisseurs placent les capacités de calcul dans l'usine elle-même. Un petit appareil Edge dans une enceinte autonome avec alimentation et refroidissement intégrés.
- Critique pour la vie humaine :** Une faible latence est critique pour ces cas d'utilisation, car ils ont un impact direct sur la santé et la sécurité humaines. Le Device Edge offre la latence la plus faible ; par conséquent, de nombreux cas d'utilisation critiques pour la vie humaine dépendent de ce modèle.

<sup>14</sup> Les villes de niveaux 2 et 3 ont souvent une population inférieure à 1 million d'habitants et disposent actuellement rarement d'un point d'échange/de peering Internet dans la ville ou d'un datacenter Hyperscale. Les exemples incluent Austin aux États-Unis ou Berlin et Milan en Europe.

<sup>15</sup> L'optimisation de la vitesse des sites Web utilise l'Edge Computing pour réduire le temps de chargement d'une page. De nombreux fournisseurs d'e-commerce connaissent des impacts négatifs sur les revenus lorsque les sites sont plus lents, et Google a observé qu'un retard de 500 millisecondes dans la réponse au niveau des pages entraînait une diminution de 20 % du trafic.

<sup>16</sup> Lorsque l'Edge se situe au niveau des sites d'accès, il se trouve sur des sites ou des points de présence appartenant à un opérateur télécom (p. ex., antennes de téléphonie cellulaire, bureaux centraux ou nœud d'un FAI). LF Edge a établi l'Edge d'accès au sein de son infrastructure de périphérie.

Dans la pratique, les entreprises tiennent compte d'autres facteurs en plus de leurs exigences relatives au cas d'utilisation lorsqu'elles prennent des décisions concernant l'infrastructure. Ces considérations importantes comprennent :

- **Environnement** : La température, la pollution et la présence de particules ont toutes un impact sur l'infrastructure requise (p. ex., le degré de refroidissement et de filtration). Le bruit produit, ceci incluant le bruit électrique, doit également être pris en compte, en particulier si l'espace sert également de bureau. Par exemple, les câbles de communication ne peuvent pas être acheminés à proximité des cages d'ascenseur.
- **Cas d'utilisation** : La quantité et la vitesse à laquelle les données doivent être traitées influencent les décisions concernant la proximité de la capacité de calcul avec le dispositif terminal. Le type de charge de travail (c.-à-d., intensive en calcul par rapport à intensive en stockage) a également un impact sur l'infrastructure Edge, car des charges de travail plus intensives en calcul (p. ex., vidéo haute définition) requièrent plus d'énergie et donc plus de refroidissement.

### La 5G va accélérer l'adoption du Edge

La 5G sera un facteur significatif dans la détermination de l'adoption du Edge, du fait le déploiement de la 5G agit comme un catalyseur pour le passage au Edge. Les régions actuellement en avance en termes de déploiement de la 5G (Amérique du Nord, Europe et Asie de l'Est) seront donc à l'avant-garde de l'adoption du Edge. Pour en savoir plus sur la manière dont les cas d'utilisation Edge bénéficieront de la 5G, consultez les [recherches précédentes de Vertiv](#).

“

La 5G commence maintenant et sa généralisation prendra 3 à 5 ans sur les grands marchés développés. Nous pensons que cela va accélérer le développement du Edge.

**VP Innovation,  
Entreprise de relais  
téléphoniques leader**

”

“

Cela constitue un défi, car ces bureaux n'ont jamais été conçus pour abriter des équipements IT, aussi nous devons nous y rendre et mettre à jour le système électrique. Maintenant, nous produisons de la chaleur dans l'espace, et nous devons donc parler de refroidissement. Particulièrement s'il s'agit d'un espace où des individus travaillent, nous ne voulons pas les exposer à un excès de chaleur, et nous ne voulons pas rendre l'ensemble trop bruyant pour eux non plus.

**Architecte solutions techniques,  
Technologie à l'échelle mondiale**

”

- **Équipement et infrastructure existants** : La décision de déployer une infrastructure Edge dans un datacenter existant plutôt que d'opter pour un nouveau déploiement autonome dépend au final de l'existence ou non d'un datacenter existant. Pour un Micro Edge, la forme spécifique de l'infrastructure est déterminée par l'espace dans lequel elle doit s'intégrer (p. ex., s'il n'y a pas suffisamment d'espace au sol, l'infrastructure doit être fixée au mur).
- **Opérations d'entreprise** : Le choix entre la mise à niveau d'un datacenter existant sur site et l'introduction d'un nouveau déploiement autonome dépend également de la capacité de l'entreprise à se permettre les temps d'arrêt requis pour améliorer son infrastructure existante. Les entreprises pour lesquelles les temps d'arrêt sont coûteux peuvent bénéficier de l'acquisition d'un datacenter préfabriqué pouvant être construit hors site et déployé rapidement.
- **Sécurité et maintenance** : Si l'infrastructure Edge se trouve dans un lieu exposé où des individus pourraient l'endommager, l'enceinte doit être conçue avec une sécurité supplémentaire. Si les employés doivent régulièrement entretenir ou interagir avec l'équipement IT, il doit être facilement accessible (p. ex., non hors de portée au plafond).
- **Infrastructure de communication** : Si l'Edge se trouve dans un emplacement éloigné et qu'une infrastructure n'existe pas pour acheminer les données sur le réseau (p. ex., exploitation minière, agriculture), une solution sur site plus robuste est nécessaire.

## Explorer les modèles d'infrastructure Edge : Recommandations clés

### Device Edge



Les cas d'utilisation qui tirent parti d'un Device Edge comprennent ceux de l'archétype critique pour la vie humaine, tels que drones, véhicules autonomes, chirurgie robotique et surveillance des patients à l'hôpital. Un Device Edge est adapté puisqu'il peut répondre aux exigences de mobilité d'un appareil, tel qu'un drone, pour naviguer de manière autonome dans le contexte de l'environnement dans lequel il se déplace. Il fournit également une latence ultra-faible, ce qui est nécessaire pour les cas d'utilisation critiques pour la vie humaine. Enfin, il permet à certains aspects du cas d'utilisation de fonctionner (p. ex., navigation, alarmes locales) même lorsque la connectivité n'est pas disponible en raison d'une couverture limitée ou d'une défaillance du réseau.

Par conséquent, le secteur de la santé est l'un des secteurs avec un taux d'adoption élevé du Device Edge, car de nombreux cas d'utilisation devront détecter des situations mettant en jeu le pronostic vital d'individus rapidement et de manière fiable, que les patients soient hospitalisés ou soignés à distance. Le secteur de la fabrication a également des cas d'utilisation critiques pour la vie humaine, c'est pourquoi les systèmes de commande machine fonctionnent principalement sur l'équipement lui-même (une forme de Device Edge).

#### Recommandations clés lors du déploiement d'un Device Edge :

- L'add-on Device Edge est plus approprié pour la modernisation d'équipements existants, mais les déploiements entièrement nouveaux peuvent envisager d'intégrer le calcul au niveau l'appareil. Cependant, il s'agit souvent d'appareils propriétaires qui ne se prêtent pas à l'intégration avec des fonctionnalités d'Edge Computing génériques.
- Le Device Edge a une capacité de calcul limitée. L'ajout de capacités de calcul supplémentaires rendra les dispositifs terminaux beaucoup plus lourds, aussi tenez toujours compte du compromis requis en termes de rapport puissance/poids<sup>17</sup> (ce qui est une question plus importante dans les cas où l'appareil est alimenté par batterie ou n'a pas accès à une alimentation électrique).
- Soyez attentif aux données collectées par le dispositif terminal. Les cas d'utilisation, tels que caméras de sécurité intelligentes, infrastructure de trafic connectée et drones, collectent des données visuelles ou de localisation concernant les personnes. Il est donc important d'être conscient des défis potentiels concernant la confidentialité et le partage des données, car cela pourrait être une question litigieuse.

<sup>17</sup> Le compromis entre puissance, poids et coût des casques AR/VR est exploré dans [Apple Glass : Un moment iPhone pour la 5G ?](#)

## Micro Edge



- Si les décisions concernant les logiciels, le matériel et l'infrastructure sont prises par différentes parties prenantes, maintenez l'alignement entre ces parties prenantes afin que les décisions soient prises en parallèle et non de manière séquentielle, car cela permet de développer une solution plus efficace.
- Sélectionnez votre type d'équipement. Un équipement renforcé est conçu pour des environnements moins contrôlés, de sorte qu'il peut supporter 50 °C. Les entreprises peuvent plutôt utiliser des serveurs génériques, vendus sur étagère (COTS<sup>18</sup>), qui sont

<sup>18</sup> COTS – Produits vendus sur étagère qui sont facilement disponibles à la vente et sont conçus pour s'intégrer facilement aux systèmes existants (plutôt que d'être sur mesure).

Un Micro Edge peut être situé à proximité de la source de données en raison de sa petite taille et de sa facilité relative de déploiement (par rapport à un datacenter de plus grande taille). Il offre donc une faible latence et réduit le coût de transmission des données, ce qui en fait un modèle d'infrastructure approprié pour les cas d'utilisation dans les trois archétypes suivants : Grand volume de données, sensibilité à la latence humaine, sensibilité à la latence machine à machine et critique pour la vie humaine. Dans les industries disposant d'un espace restreint, telles que le secteur retail ou de l'éducation, un Micro Edge constitue une solution attrayante, car elle limite l'immobilier requis, permettant ainsi le déploiement de capacités informatiques avec un encombrement réduit. Par exemple, une grande chaîne de supermarchés avec 16 000 sites en Europe déploie un Micro Edge dans les magasins pour la collecte et le traitement de données locales, et ajoute également des datacenters centraux pour l'agrégation et la gestion IT générale.

### Recommandations clés lors du déploiement d'un Micro Edge :

- Considérez l'espace disponible (il peut être nécessaire de fixer aux murs ou au plafond), la fonction de l'espace (si des clients ou des employés sont présents) et les exigences de sécurité (lorsque l'infrastructure est facilement accessible, une couche physique de sécurité est nécessaire). Les déploiements de Micro Edge couvrent souvent des zones avec différents types d'alimentations électriques, divers règlements, différents types d'accès au site (p. ex., hauteur de l'ascenseur), de formes de contrôle du site (responsable de magasin, responsable d'usine) ainsi que de différents types et niveaux d'expertise technique.



moins chers, mais la durée de conservation de ces serveurs est considérablement réduite lorsqu'ils fonctionnent au-dessus de 30 °C. Alors que les deux types de matériel nécessitent une enceinte, l'infrastructure de support pour les serveurs COTS doit offrir un meilleur contrôle de la température, de l'humidité et de l'alimentation. Un équilibre économique entre standardisation et personnalisation spécifique au site est nécessaire.

## Datacenter Edge distribué



Comme un Micro Edge, les datacenters Edge distribués sont situés sur le site de l'entreprise et conviennent à de nombreux cas d'utilisation sectoriels, car ils offrent une faible latence et des coûts de bande passante réduits. Les recherches ont révélé que les opérateurs télécoms utilisent des datacenters Edge distribués pour héberger à la fois des applications grand public et leurs propres fonctions réseau internes, qui sont sensibles à la latence machine à machine. De même, les entreprises de fabrication de taille moyenne et grande utiliseront ces datacenters pour leurs cas d'utilisation de l'Internet des objets (Internet of Things, IOT). Pour les sites de fabrication de taille moyenne, la plupart de l'infrastructure Edge résidera dans un datacenter à huit racks.

### Recommandations clés lors du déploiement d'un datacenter Edge distribué :

- Un investissement peut être nécessaire pour améliorer un datacenter existant ou une salle réseau existante, et le temps de déploiement des modifications pourrait avoir un impact coûteux sur les opérations. Cet impact sur les temps d'arrêt doit être comparé au coût d'achat d'un nouveau datacenter préfabriqué pouvant être déployé rapidement sur site.
- Il est recommandé d'intégrer des capacités de recharge dans le datacenter afin de maintenir la flexibilité à l'avenir, pour autant il convient de noter qu'une « surconstruction » pour se préparer à tous les résultats est coûteuse et peut ne pas être nécessaire. Pour trouver l'équilibre entre ce qui est nécessaire aujourd'hui et ce qui sera nécessaire demain, les utilisateurs doivent envisager l'évolution de leur cas d'utilisation Edge dans leur secteur donné.
- Lors du développement de capacités de redondance au sein du datacenter, tenez compte à la fois de la valeur des applications exécutées et de la stabilité de l'environnement (p. ex., dans certains pays, le réseau n'est pas fiable, aussi le risque de coupures d'alimentation est important).
- Parfois, il n'est pas nécessaire de déployer un datacenter Edge distribué sur le site de l'entreprise, car un déploiement à proximité des locaux répond à des exigences telles que celles liées à la latence ou à la sécurité. Il pourra s'agir d'un site appartenant à l'entreprise ou d'un site multi-locataires desservant plusieurs entreprises.
- Si un datacenter Edge distribué est utilisé comme site de colocation, il doit disposer de couches de sécurité et d'isolation afin de fournir cette expérience Edge Computing multi-locataires. Cela peut inclure des portes, verrous et caméras.

## Datacenter Edge régional



Un datacenter Edge régional agit soit comme un site Edge Computing, soit comme un site intermédiaire où les données Edge sont envoyées aux fins de prétraitement avant d'être envoyées au cloud. Il satisfait à la fois aux cas d'utilisation à faible latence et à forte consommation de données, aussi tous les archétypes Edge tirent parti des datacenters Edge régionaux. Les cas d'utilisation sensibles à la latence humaine, en particulier, s'appuient sur les datacenters Edge régionaux, car les déploiements Edge sur site (p. ex., streaming multimédia à faible latence ou jeux immersifs) ne constituent pas une option.

Les datacenters Edge régionaux sont souvent adoptés dans le secteur retail, car ils peuvent réduire le besoin de déployer une infrastructure informatique dans les magasins individuels.<sup>19</sup> Dans les cas où le retailer a investi dans des déploiements individuels sur site, le datacenter peut agir comme un site intermédiaire de traitement des données.

### Recommandations clés lors du déploiement d'un datacenter Edge régional :

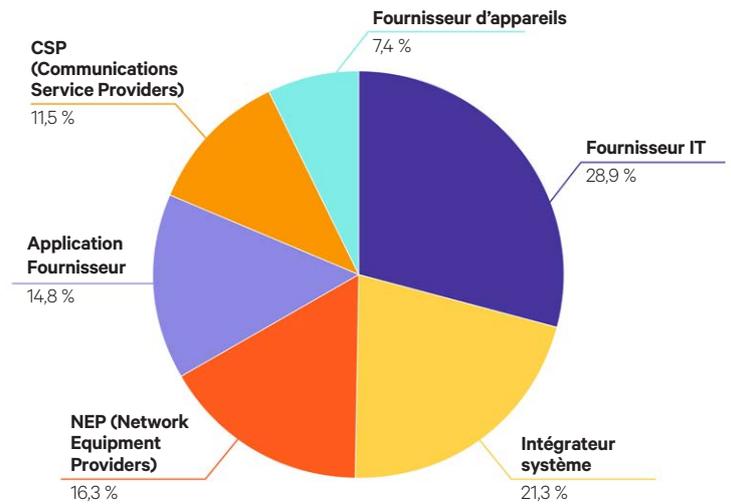
- La sécurité et l'isolation sont des nécessités (car de nombreux datacenters régionaux sont des sites multi-locataires). Les clients doivent s'assurer que le datacenter dispose de mécanismes adéquats pour sécuriser l'infrastructure et les données des locataires.
- Tenez compte du cas d'utilisation spécifique lors de la conception de l'infrastructure d'Edge Computing (p. ex., des charges de travail plus gourmandes en calcul nécessiteront probablement plus d'énergie et, par conséquent, plus de refroidissement).
- L'emplacement est une considération clé. Si la souveraineté des données est un facteur, les données peuvent devoir être stockées dans la juridiction des clients finaux. Toutefois, si le facteur clé est la latence (< 50 millisecondes), ciblez un emplacement stratégiquement important qui réduit la latence sur autant de sites finaux que possible. Il s'agira souvent d'un datacenter situé au niveau ou très proche d'un échange Internet majeur.
- Les principaux fournisseurs de cloud public étendent leur cloud public aux datacenters locaux (par exemple, AWS Local Zones), ce qui permettra aux entreprises de distribuer leurs applications cloud plus facilement. Cependant, il existe deux considérations clés : les fournisseurs de cloud public sont à un stade précoce du déploiement de ces clouds locaux, et certaines applications (ainsi que les données) ne seront pas adaptées au stockage et au traitement sur un cloud public (en partie en raison des règlements de services publics).

<sup>19</sup> Selon le [Worldwide Edge Spending Guide de International Data Corporation \(IDC\)](#), le secteur retail est le deuxième secteur le plus important et connaissant la croissance la plus rapide sur le marché européen de l'Edge d'entreprise.

## Un écosystème est requis pour développer l'Edge

L'infrastructure n'est qu'une pièce du puzzle pour toute organisation cherchant à mettre en œuvre des solutions basées sur l'Edge. Il existe de nombreux éléments qui influencent la construction du Edge : logiciel, matériel, infrastructure, orchestration, gestion, etc., et les entreprises auront du mal à coordonner l'ensemble de ces éléments par elles-mêmes.

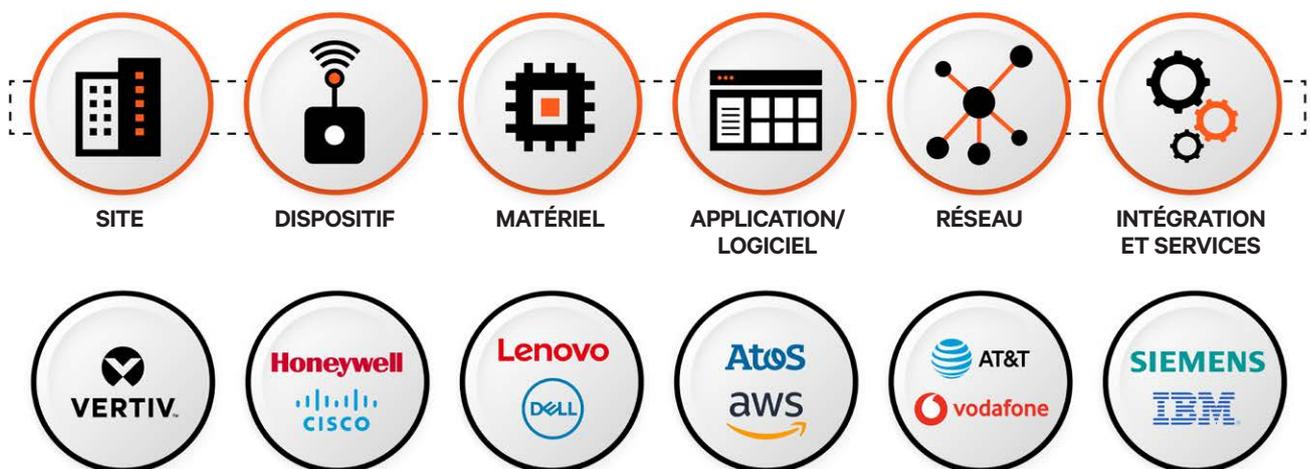
Il n'existe pas de solution unique ou universelle concernant la manière dont les entreprises souhaitent acheter ces solutions TIC (technologies de l'information et de la communication) et Edge de nouvelle génération. Les recherches montrent que 34 % des entreprises préfèrent une approche de type « do-it-yourself », suivant laquelle elles sélectionnent différents composants auprès de chaque fournisseur. Les 66 % qui préfèrent que des solutions complètes proviennent d'un fournisseur principal, varient en termes de qui sera ce fournisseur principal : Fournisseur IT versus intégrateur de systèmes versus fournisseur d'équipement réseau, etc.



Source : Enquête STL Partners auprès de 699 professionnels de l'industrie à l'échelle mondiale, mai 2021

Qu'il s'agisse d'une entreprise développant son propre Edge ou d'un hébergeur déployant une infrastructure Edge pour exécuter des applications ou permettre à d'autres d'exécuter des charges de travail, la collaboration avec d'autres acteurs de l'écosystème de Edge Computing est essentielle à la réussite. Établir des relations solides avec des spécialistes de l'industrie (par exemple, Siemens ou Honeywell dans le domaine de la fabrication) garantit que les solutions répondent aux besoins spécifiques du secteur et peuvent s'intégrer avec succès aux infrastructures et systèmes existants.

### La chaîne de valeur du Edge Computing :



## Conclusions et Recommandations

L'infrastructure Edge reste un sujet compliqué (comme le démontrent les entretiens avec des professionnels de l'industrie). Néanmoins, le cadre du modèle d'infrastructure Edge défini dans ce rapport peut aider les entreprises à naviguer à travers l'éventail des solutions Edge disponibles et disposer de conseils concernant les choix d'infrastructure appropriés.

En regardant au-delà des modèles d'infrastructure Edge, Vertiv reconnaît qu'il existe des complexités associées à la tâche pratique de construction d'une infrastructure Edge qui sont uniques à chaque entreprise. Un outil Web interactif a été développé en vue de permettre aux entreprises et autres opérateurs de datacenter d'explorer en profondeur les cas d'utilisation clés. Les entreprises seront en mesure de mieux comprendre la charge de travail associée et les caractéristiques de l'infrastructure, et ainsi d'éclairer leurs décisions concernant la conception, la construction et les déploiements d'infrastructure.

### Les autres recommandations clés comprennent :

#### Entreprises

- **Identifier un cas d'utilisation d'ancrage.** La nature des cas d'utilisation futurs et des besoins associés en matière d'Edge Computing demeure incertaine. Le premier cas d'utilisation doit justifier l'analyse d'entreprise d'une construction initiale, assurez-vous donc de disposer d'une bonne compréhension de la raison pour laquelle votre cas d'utilisation requiert l'Edge Computing. Comprendre quelles sont les caractéristiques de charge de travail constituant le facteur clé du déploiement Edge (p. ex., latence, bande passante, sécurité) aidera également à éclairer les décisions concernant l'infrastructure.
- **Être prêt à gérer une variété de modèles d'infrastructure Edge.** Par exemple, de nombreux retailers optent pour un Micro Edge dans leurs magasins et le complètent avec un datacenter Edge distribué à proximité des magasins pouvant filtrer et agréger les données provenant de différents emplacements, en envoyant uniquement les informations nécessaires au cloud.
- **Essayer de ne pas définir un modèle unique pour tous les scénarios.** Même au sein des types de modèles, il existera des variations étant donné que différents emplacements ont des environnements hérités différents. Les entreprises ayant une couverture multinationale seront confrontées à des différences géographiques en termes de climat, de pollution, d'alimentation électrique, de réglementation, etc. (p. ex., l'UE réglemente le nombre de décibels autorisés ce qui pourrait limiter l'infrastructure en matière de ventilateurs ou de dispositifs de commutation).

#### Fournisseurs de solutions

- **Infrastructure Edge à l'épreuve du temps.** Comprendre les cas d'utilisation que les clients adoptent aujourd'hui et prévoient d'adopter à l'avenir, et mettre en place des capacités de réserve (stockage, calcul, etc.) le cas échéant. L'adoption de modèles de déploiement plus flexibles réduira les risques.
- **Collaborer avec l'écosystème.** Edge n'est pas un produit unique à vendre par un fournisseur unique, mais une solution que plusieurs acteurs de l'écosystème développent ensemble. Les solutions doivent donc être standardisées, de sorte qu'il soit facile pour les clients d'utiliser une solution en tant que composant. Le partenariat est également important, en particulier lorsque l'on cherche à répondre à des besoins hautement spécifiques à un secteur.
- **Considérer de nouveaux modèles économiques.** Répliquer ce qui a été fait avec le cloud n'est pas possible. L'infrastructure Edge a des besoins spécifiques, il est donc important de prendre en compte les modèles économiques qui garantissent que l'alimentation, le refroidissement, la sécurité et l'espace sont optimisés avec de nouvelles façons de garantir des économies d'échelle.



# Annexe : Glossaire

<b>ACCÉDER AU EDGE</b>	Un site Edge au sein du réseau de télécommunications qui connecte les abonnés au réseau principal de l'opérateur principal, puis à d'autres réseaux, à Internet et aux clouds Hyperscale.
<b>SITE OU SERVICE DE COLOCATION</b>	Un site de colocation ou « colo » est un site de datacenter dans lequel une entreprise peut louer des espaces pour des serveurs et d'autres matériels informatiques. Généralement, un site colo fournit le bâtiment, le refroidissement, l'alimentation, la connectivité à d'autres ou à Internet, et la sécurité physique, tandis que le client fournit serveurs et stockage.
<b>ENVIRONNEMENT CLIMATISÉ, CONTRÔLÉ</b>	Environnements avec des systèmes dédiés en place pour contrôler divers facteurs, notamment la température et l'humidité, les particules de poussière, la pollution, etc.
<b>DATACENTER</b>	Un site physique que les organisations utilisent pour héberger leurs applications et données critiques. La conception d'un datacenter est basée sur un réseau de ressources informatiques et de stockage permettant la livraison d'applications et de données partagées. Les composants clés de la conception d'un datacenter comprennent les routeurs, les commutateurs, les pare-feu, les systèmes de stockage, les serveurs et les contrôleurs de livraison d'applications.
<b>EDGE COMPUTING</b>	Cette infrastructure informatique physique est positionnée entre l'appareil et le cloud Hyperscale, prenant en charge diverses charges de travail. L'Edge Computing rapproche les fonctionnalités de traitement de l'utilisateur final/de l'appareil final/de la source de données, ce qui élimine la distance aux datacenters des fournisseurs de cloud et réduit ainsi la latence.
<b>FORMAT</b>	Conception globale et fonctionnalité des systèmes matériels.
<b>HYPERSCALE</b>	En informatique, l'Hyperscale est la capacité à atteindre une échelle massive, en particulier pour le Big Data et le Cloud Computing. Aujourd'hui, AWS, Azure et Google Cloud sont considérés comme des « Hyperscalers ».
<b>Armoire IT/RÉSEAU</b>	Une armoire ou une salle pièce où le câblage électrique et le matériel de réseau informatique sont installés.
<b>EDGE COMPUTING MULTI-ACCÈS (MEC)</b>	Type d'architecture réseau fournissant des fonctionnalités de cloud computing et un environnement de service IT à la périphérie du réseau.
<b>SUR SITE</b>	Également désignée « sur site » (on premises ; on-prem), il s'agit de la technologie hébergée dans les limites physiques du propre site d'une entreprise.
<b>MAINTENANCE PRÉDICTIVE</b>	Processus de surveillance des données provenant des capteurs d'équipement afin de s'assurer qu'il est en bonne condition et pour signaler de manière préventive s'il est nécessaire d'effectuer des réparations, éliminant éventuellement le besoin d'une maintenance planifiée.
<b>MATÉRIEL RENFORCÉ</b>	Matériel conçu spécifiquement pour résister aux environnements difficiles tels que : pollution extérieure, températures élevées ou basses, humidité, etc.
<b>AUTONOME</b>	Capable de fonctionner indépendamment d'autres matériels ou logiciels.
<b>STATION TÉLÉCOMS DE BASE</b>	Station de transmission et de réception dans un emplacement fixe, comprenant une ou plusieurs antennes de réception/transmission, une antenne parabolique et des circuits électroniques, utilisés pour gérer le trafic cellulaire.



 **PARTNERS** Ce rapport de recherche a été développé avec le soutien de STL Partners.

**Vertiv.fr | Vertiv France SAS, Bâtiment Liège, 1 Place des États-Unis 94150, Rungis, France**

© 2021 Vertiv Group Corp. Tous droits réservés. Vertiv™ et le logo Vertiv sont des marques déposées ou commerciales de Vertiv Group Corp. Tous les autres noms et logos mentionnés sont des noms de produits, des marques commerciales ou déposées qui appartiennent à leurs propriétaires respectifs. Même si toutes les précautions ont été prises pour garantir l'exactitude et l'exhaustivité des informations figurant dans le présent document, Vertiv Group Corp. ne saurait être tenu responsable et décline toute responsabilité en cas de dommages résultant d'une utilisation de ces informations ou de quelconques erreurs ou omissions. Les descriptifs techniques, remises et autres offres promotionnelles sont susceptibles d'être modifiés à la seule discrétion de Vertiv après notification.