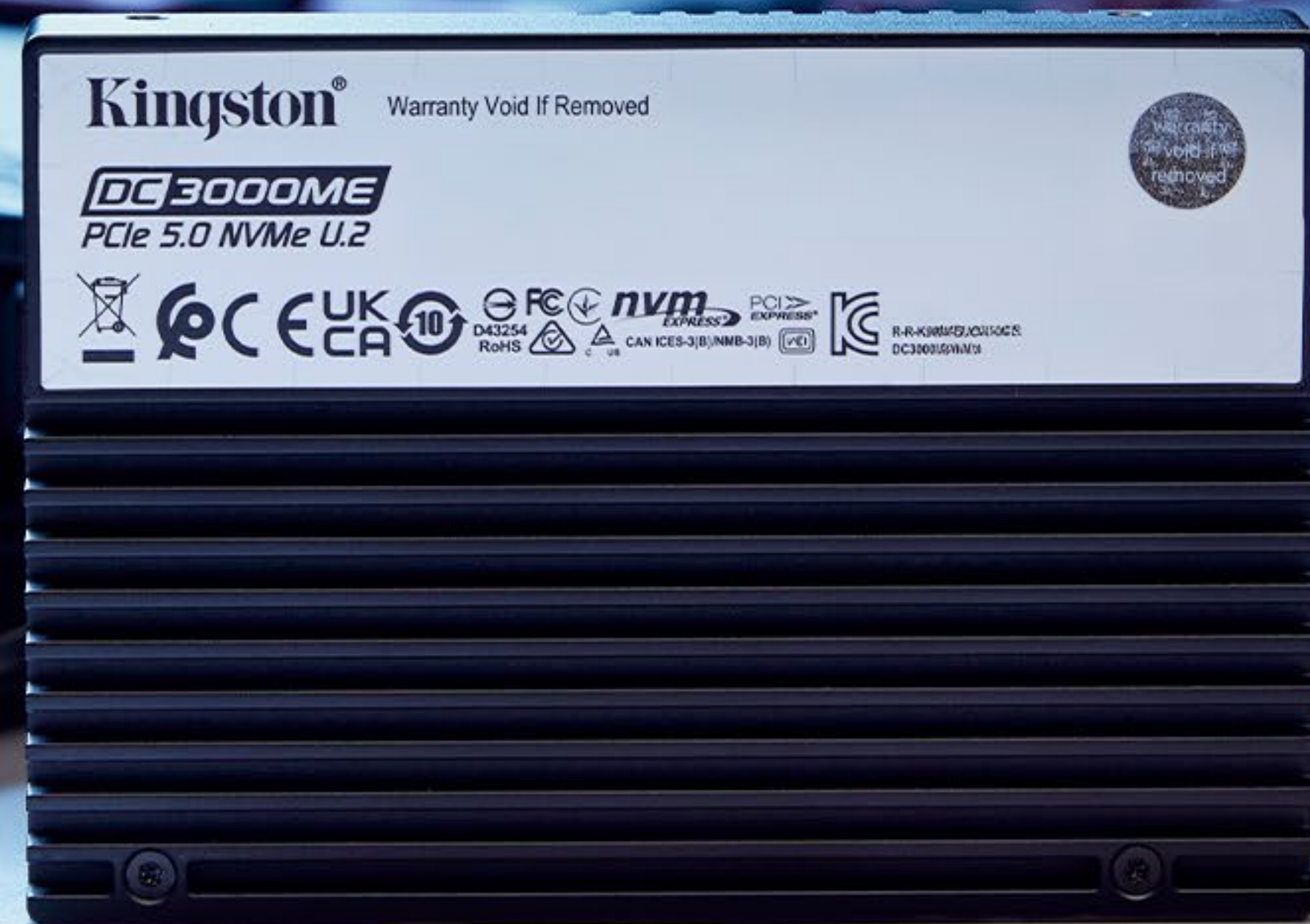


OPTIMISER LE STOCKAGE DES DATACENTERS AVEC DES SSD PCIe GEN5 POUR L'IA ET LE BIG DATA



Avant-propos et contenu

Dans un monde où les données sont reines, les infrastructures qui les soutiennent sont au cœur de toutes les attentions. De l'IA générative à l'analyse en temps réel, la demande en matière de vitesse, de fiabilité et d'efficacité s'accélère, tout comme le besoin en stockage pour datacenters de nouvelle génération. Aussi, les SSD, en particulier ceux qui utilisent la technologie PCIe Gen5, sont devenus des piliers essentiels pour les environnements informatiques modernes.

Qu'est-ce qui motive cette tendance ? Comment les goulots d'étranglement affectent-ils l'utilisation du GPU ? Pourquoi un stockage d'entreprise économe en énergie est-il aussi important que les IOPS ? Et comment les responsables informatiques peuvent-ils protéger l'intégrité des données tout en préparant leur infrastructure pour l'avenir ?

Cet eBook répond à ces questions et explique comment améliorer les performances du stockage pour datacenters, ainsi que et les principales tendances du secteur, du coût croissant des temps d'arrêt à la promesse d'une infrastructure durable. Grâce aux connaissances des experts en stockage Flash de Kingston, découvrez comment les SSD avancés transforment les datacenters pour l'ère de l'IA et au-delà.

Sommaire	Pages
Contributeurs	3
Augmentation massive des données	4
Goulots d'étranglement au niveau des performances et importance de la vitesse	5
Coût des temps d'arrêt et de la latence	6
Fiabilité et intégrité des données	7
Gérer la consommation d'énergie et la durabilité	8
Préparer votre infrastructure de datacenter pour l'avenir	9
Le rôle de Kingston dans l'optimisation des datacenters	10-12
Résumé et informations sur Kingston	13



Contributeurs

Cet eBook a été créé par deux experts Kingston en technologies de stockage.



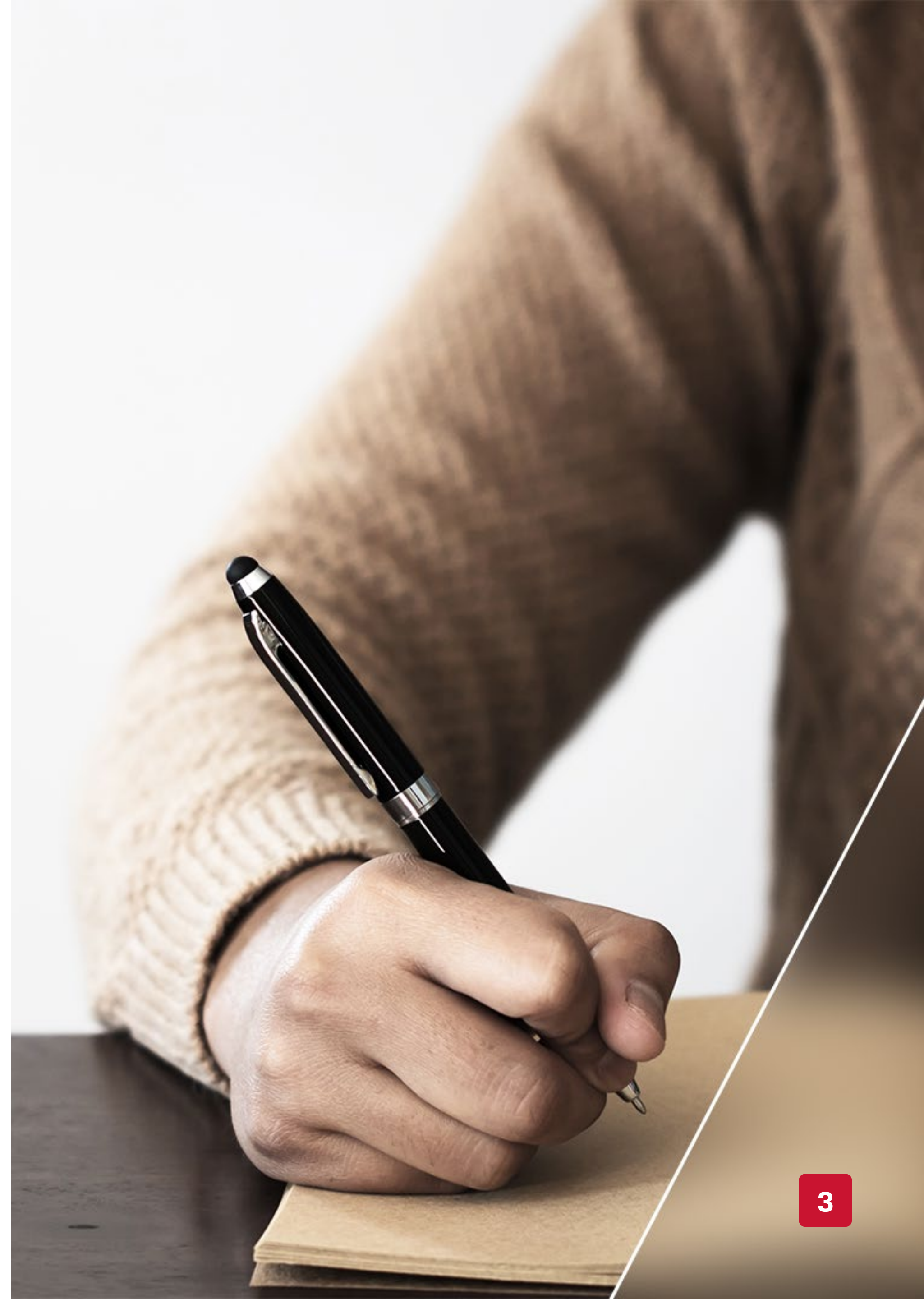
Louis Kaneshiro | Kingston Technology

Louis Kaneshiro est Directeur de la technologie chez Kingston Technology. Au cours de ses 30 années chez Kingston, dont les 15 dernières consacrées aux SSD, il a dirigé le groupe des ressources technologiques, qui est aujourd'hui une équipe mondiale, avant de lancer le département d'ingénierie des produits SSD (SPED). Avant d'occuper son poste actuel, Louis a travaillé dans le support technique et comme ingénieur d'application terrain pour la division OEM de Kingston.



Tony Hollingsbee | Kingston Technology Europe

Tony travaille chez Kingston Technology depuis 23 années, au cours desquelles où il a occupé divers postes dans la vente, notamment celui de directeur des ventes internes pour les gammes mémoires Flash et SSD de l'entreprise. Depuis des 12 dernières années, il occupe le poste de Directeur commercial SSD pour la région EMEA, et travaille en étroite collaboration avec les équipes commerciales et marketing afin de soutenir le développement commercial et de stimuler la croissance régionale.



Augmentation massive des données

La quantité exponentielle de données créées et consommées dans le monde devrait atteindre plus de 394 zettaoctets d'ici 2028¹. Cette explosion des volumes de données est en grande partie due aux progrès en matière d'intelligence artificielle (IA) et d'Internet des objets (IoT), ainsi qu'à l'augmentation de la consommation de médias numériques. Les applications d'IA et de machine learning génèrent et consomment d'énormes ensembles de données. Parallèlement, les technologies émergentes telles que les véhicules autonomes et les villes intelligentes stimulent la génération de données en temps réel à une échelle sans précédent.

Ces tendances nécessitent des systèmes de stockage à faible latence et haut débit. Cette demande en stockage évolutif et sécurisé est également accrue par les charges de travail des entreprises, les applications cloud natives et la conformité aux réglementations. Pour faire face à cette augmentation, les datacenters font évoluer leurs infrastructures de stockage. Le format standard pour les entreprises et les datacenters (EDSFF) gagne du terrain. Et des technologies de stockage innovantes, comme les solutions d'archivage à base de céramique avec des capacités de conservation des données supérieures à 5 000 ans, sont développées pour répondre aux exigences de conservation des données à long terme.

Bien que le paysage actuel des formats de stockage pour datacenters soit vaste et varié, les formats traditionnels comme les SSD U.2, connus pour leur interface PCIe et leur adaptabilité aux environnements axés sur les performances, ont été largement adoptés. C'est particulièrement vrai chez les fabricants OEM de serveurs et les fabricants de châssis de serveurs. Ces avancées visent toutes à répondre aux besoins croissants des datacenters modernes en matière de stockage, face à cette prolifération incessante des données.



Les SSD tels que le DC3000ME jouent un rôle essentiel pour répondre aux nouvelles tendances informatiques, en particulier dans les domaines de l'IA, du machine learning et de l'Edge Computing. Notre design anticipe le besoin en solutions de stockage offrant une bande passante plus élevée, une latence plus faible et une meilleure efficacité énergétique.

Tony Hollingsbee | Kingston Technology Europe



Alors que les SSD E1.x et E3.x gagnent en popularité, le format U.2 reste de loin le format dominant, avec plus de 60 % des pétaoctets stockés dans des serveurs résidant dans un SSD U.2.

Louis Kaneshiro | Kingston Technology



1. Volume de données/informations créées, capturées, copiées et consommées dans le monde entre 2010 et 2023, avec des prévisions pour 2024 à 2028 (en zettaoctets). <https://www.statista.com/statistics/871513/worldwide-data-created/>



Avec l'essor de l'IA, les performances du stockage sont plus importantes que jamais pour optimiser l'utilisation des GPU.

Louis Kaneshiro | Kingston Technology



Goulets d'étranglement au niveau des performances et importance de la vitesse

Avec l'explosion des volumes de données et la complexité croissante des applications, les responsables de datacenters sont confrontés à des défis de plus en plus importants. L'un des problèmes les plus courants est la variabilité des performances. De nombreuses solutions de stockage pour datacenters d'entreprise ne parviennent pas à fournir un débit et une latence constants, ce qui entraîne des inefficacités dans les performances des applications et l'utilisation des ressources.

La complexité des déploiements complique encore davantage la situation, nécessitant souvent des configurations personnalisées, des réglages des firmwares et une orchestration complexe pour répondre aux besoins spécifiques des charges de travail. À cela vient s'ajouter le problème de la longévité des SSD. En particulier, les charges de travail à écriture intense, comme la journalisation, la mise en cache et la formation de l'IA, exercent une pression énorme sur les SSD, ce qui accélère leur usure et augmente le risque de temps d'arrêt.

Les charges de travail IA dépendent fortement des GPU pour le calcul, mais si le stockage sous-jacent ne suit pas, l'utilisation des GPU chute, ce qui empêche de tirer parti des investissements et d'atteindre les objectifs de performance. Les SSD hautes performances pour les charges de travail IA, avec leur latence ultra-faible et leur débit élevé, sont essentiels pour soutenir les pipelines de données qui alimentent les GPU en temps réel.

L'autre défi majeur des datacenters est la vitesse. Les charges de travail modernes, en particulier l'IA, le machine learning et l'analyse en temps réel, exigent un accès instantané à de grandes quantités de données. Toute latence au niveau de la couche de stockage peut bloquer le traitement, retarder les informations et réduire l'efficacité globale du système. Le stockage à haute vitesse n'est plus un luxe, c'est une nécessité pour les ressources informatiques hautes performances comme les GPU et les TPU.

Coût des temps d'arrêt et de la latence

Pour les responsables informatiques, les temps d'arrêt et la latence sont également des préoccupations majeures, car ils peuvent entraîner des pertes financières importantes et des perturbations opérationnelles. De récentes études montrent que le coût d'une heure de temps d'arrêt dépasse les 300 000 \$ pour 90 % des entreprises, 41 % d'entre elles estimant que ce coût se situe entre 1 et 5 millions\$².

La latence aggrave encore ces difficultés en ralentissant l'accès aux données, ce qui est particulièrement préjudiciable dans les environnements à forte intensité de données. Les retards d'accès aux données dégradent l'expérience utilisateur, réduisent la productivité et entraînent des pertes de revenus. Dans le trading à haute fréquence, par exemple, la latence est un facteur critique où quelques millisecondes peuvent faire la différence entre des millions de bénéfices ou de pertes.

Les SSD avancés offrent une solution efficace à ces problèmes. Grâce à des technologies telles que les interfaces PCIe Gen5 et les architectures 3D NAND avancées, les SSD modernes offrent des vitesses de lecture et d'écriture nettement plus élevées que les disques durs traditionnels (HDD) et les SSD SATA.

Ces performances accrues accélèrent non seulement l'accès aux données, mais améliorent également la fiabilité du système, réduisant ainsi la probabilité et l'impact des temps d'arrêt. En intégrant ces SSD avancés à leur infrastructure, les entreprises peuvent gagner en efficacité opérationnelle et en résilience.



Les SSD de nouvelle génération pour l'Edge Computing assurent un traitement des données à faible latence et à haut débit qui prend en charge les opérations en temps réel essentielles aux systèmes distribués.

Louis Kaneshiro | Kingston Technology



2. Le coût d'une heure de temps d'arrêt dépasse les 300 000 \$ pour 90 % des entreprises, 41 % d'entre elles déclarant que ce coût est compris entre 1 et 5 millions \$. <https://itic-corp.com/itic-2024-hourly-cost-of-downtime-report/>

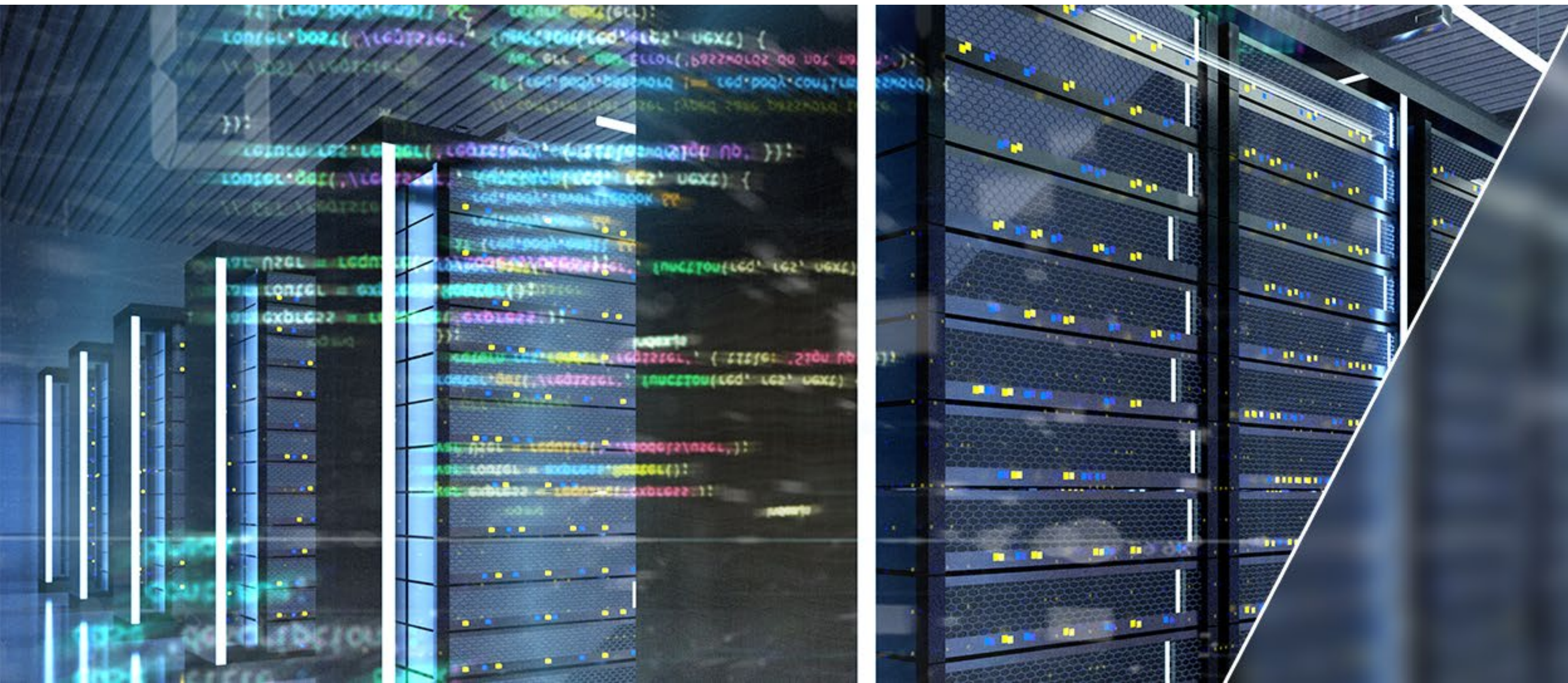


Fiabilité et intégrité des données

Dans l'économie numérique actuelle, la fiabilité et l'intégrité des données sont indispensables pour les datacenters qui gèrent des charges de travail critiques. Alors que les entreprises s'appuient de plus en plus sur l'analyse en temps réel, les applications cloud natives et les systèmes basés sur l'IA, même une perte ou une corruption mineure de données peut entraîner des revers opérationnels et nuire à la réputation.

La fiabilité des données dépend fortement de la robustesse de l'infrastructure de stockage. L'endurance des lecteurs, les mécanismes de correction des erreurs et les capacités de basculement sont autant d'éléments essentiels pour maintenir une disponibilité et une intégrité des données constantes. Les supports de stockage doivent non seulement résister à l'usure physique, mais aussi détecter et corriger les erreurs au niveau des bits avant qu'elles n'affectent les applications.

Les SSD de classe entreprise, dotés de technologies telles que la protection des données de bout en bout, la protection contre les coupures d'alimentation et la correction avancée des erreurs (comme le LDPC), sont conçus pour protéger les données même dans des conditions de charge extrêmes. Et avec la prolifération des rançongiciels et des menaces de cybersécurité, il est tout aussi essentiel de préserver l'intégrité des données lors des attaques. Les SSD de nouvelle génération, conçus pour offrir endurance, cohérence et résilience, constituent la base de confiance dont les datacenters modernes ont besoin. En donnant la priorité à la fiabilité et à l'intégrité des données, les entreprises peuvent protéger leur atout le plus précieux : leurs données.



« Face à l'augmentation des besoins informatiques, l'évolutivité des SSD Gen5 permet aux datacenters de s'adapter et de répondre aux demandes de demain. Des fonctionnalités telles que la protection contre les coupures d'alimentation assistée par le matériel garantissent l'intégrité des données, ce qui est essentiel dans les applications critiques où la préservation des données est une priorité.

Louis Kaneshiro | Kingston Technology



Gérer la consommation d'énergie et la durabilité

Outre la fiabilité et l'intégrité des données, la réduction de la consommation énergétique des datacenters est une priorité absolue pour les responsables informatiques. À mesure que les objectifs de durabilité et d'efficacité opérationnelle convergent et que les volumes de données mondiaux explosent, l'énergie nécessaire pour traiter, stocker et transférer ces données connaît une augmentation sans précédent.

En 2023, la demande mondiale des datacenters s'est élevée à 340 TWh, soit environ 1,3 % de la consommation électrique mondiale. Plusieurs sources prévoient que la demande va exploser d'ici 2030, ce qui va mettre encore plus de pression sur les infrastructures électriques et les efforts de développement durable³.

Réduire la consommation d'énergie ne se limite pas au respect des réglementations environnementales et des objectifs ESG (environnementaux, sociaux et de gouvernance) des entreprises. C'est essentiel pour réduire les coûts énergétiques et garantir la viabilité d'un datacenter. Par conséquent, l'optimisation du stockage est un domaine prioritaire pour les responsables de datacenters actuels.

Les disques durs traditionnels consomment plus d'énergie et génèrent plus de chaleur que les SSD SATA. En revanche, bien que les SSD PCIe NVMe de nouvelle génération soient plus proches des disques durs en termes de consommation électrique et de dégagement de chaleur, ils offrent une efficacité énergétique nettement supérieure en termes de performances par watt. Grâce au nombre impressionnant d'IOPS que les SSD PCIe NVMe peuvent traiter, les données peuvent être rapidement accessibles et traitées par les systèmes hôtes, ce qui réduit les temps d'opération et a un impact positif sur la consommation énergétique globale.

Les SSD PCIe NVMe haute capacité permettent également un stockage plus dense dans des espaces réduits. Non seulement cela favorise des opérations plus écologiques, mais cela améliore également l'évolutivité de l'infrastructure de données. Les SSD avancés sont donc essentiels à la conception de datacenters durables et performants.

L'efficacité énergétique des SSD est particulièrement utile, car ils maintiennent des performances élevées tout en réduisant la consommation d'énergie et l'empreinte carbone globale, un facteur essentiel pour les appareils destinés aux déploiements de datacenters de milieu ou haut de gamme.

Louis Kaneshiro | Kingston Technology

3. La croissance de la demande mondiale en énergie a presque doublé en 2024 par rapport à sa moyenne récente.
<https://www.iea.org/news/growth-in-global-energy-demand-surged-in-2024-to-almost-twice-its-recent-average>

Préparer votre infrastructure de datacenter pour l'avenir

La complexité accrue des charges de travail et l'augmentation des volumes de données nécessitent des technologies évolutives et performantes. Dans ce domaine, les SSD PCIe 5.0 changent la donne, en offrant une bande passante deux fois supérieure à celle de la génération 4 et en réduisant considérablement la latence. Ils permettent un accès plus rapide aux données, améliorent l'utilisation des GPU pour les charges de travail IA et prennent en charge des déploiements de stockage plus denses.

L'intégration de SSD et d'architectures de serveurs compatibles Gen 5 permettra à l'infrastructure de répondre aux demandes futures, qu'il s'agisse d'analyse en temps réel, de machine learning ou d'Edge Computing, sans nécessiter de mises à niveau matérielles constantes et coûteuses.

Lorsqu'elles envisagent d'adopter des SSD Gen 5 pour leurs datacenters, les entreprises doivent commencer par évaluer leurs applications et leurs besoins en stockage afin de déterminer si ce niveau de performances est nécessaire. Ce processus doit débuter par une évaluation des exigences de performances pour les une, deux, cinq prochaines années par l'architecte du stockage.

Il convient de prendre en compte les différents environnements de datacenter et les prévisions de croissance du stockage (augmentation ou diminution), ainsi que la mesure et l'évaluation du retour sur investissement. À cette fin, la télémétrie des applications peut aider à recueillir des mesures de performance et obtenir des prévisions de croissance de la capacité de stockage avec une grande précision.

Il faut prendre en compte l'infrastructure matérielle actuelle, ainsi que la compatibilité du matériel actuel avec l'infrastructure PCIe 5.0 existante pour comprendre les besoins en alimentation électrique et valider les solutions de gestion thermique.



Le coût total de possession doit être évalué non seulement en prenant en compte le prix d'achat initial, mais aussi les performances par watt, les mesures de fiabilité et l'efficacité opérationnelle globale.

Tony Hollingsbee | Kingston Technology Europe



En examinant ces facteurs en profondeur et en envisageant des stratégies pérennes/en évitant les pièges, les entreprises peuvent prendre des décisions éclairées qui leur permettront d'être correctement équipées pour faire face aux difficultés d'aujourd'hui et aux avancées technologiques de demain.

Louis Kaneshiro | Kingston Technology





Le rôle de Kingston dans l'optimisation des datacenters

Les SSD tels que le [DC3000ME de Kingston](#) sont une composante essentielle pour améliorer les performances des serveurs et des GPU. Ils permettent de réduire les temps d'arrêt et de gérer facilement l'Edge Computing, l'infrastructure IA et les processus de calcul intensifs.

Représentant un énorme bond en avant dans les performances de stockage des datacenters, le DC3000ME est spécialement conçu pour répondre aux demandes exponentielles de l'IA de nouvelle génération et des charges de travail informatiques hautes performances.

Nos principales innovations techniques comprennent une interface PCIe 5.0 révolutionnaire avec une technologie de contrôleur avancée. Cette fonctionnalité offre des vitesses de lecture/écriture sans précédent, une gestion thermique intégrée et un format optimisé pour les environnements de serveurs denses. Ces avancées soulignent notre engagement à façonner un écosystème technologique robuste et agile dont bénéficieront les futures générations de datacenters. Voici comment :

Conçu pour les charges de travail intensives

Pour les charges de travail d'IA et de ML, le DC3000ME avec son interface Gen5 offre un débit de lecture et des IOPS en lecture supérieurs. Cela signifie que même un seul lecteur DC3000ME peut fournir efficacement des E/S à plusieurs GPU afin de saturer efficacement les performances. Il en résulte des temps de formation plus courts et une réduction du nombre de SSD nécessaires, en particulier par rapport à la Gen4, pour suivre le rythme de plusieurs GPU. Cela permet de réduire les coûts d'infrastructure tout en maintenant les performances.

Le DC3000ME complète l'infrastructure IA grâce à sa bande passante élevée, qui lui permet de gérer facilement les charges de travail gourmandes en données et d'utiliser les GPU en toute

fluidité. La capacité à servir plusieurs GPU simultanément accélère les activités de formation et d'inférence IA, ce qui les rend plus efficaces et plus productives. Face à l'augmentation des besoins informatiques, l'évolutivité des SSD Gen5 permet aux datacenters de s'adapter et de répondre aux demandes de demain.



Les performances du DC3000ME sont exceptionnelles, avec des vitesses de lecture séquentielle allant jusqu'à 14 000 Mo/s et des vitesses d'écriture allant jusqu'à 10 000 Mo/s. Il offre différents niveaux de puissance, ce qui assure la flexibilité en matière de consommation d'énergie et des performances évolutives.

Tony Hollingsbee | Kingston Technology Europe



Grâce à la mise en œuvre de la dernière technologie PCIe NVMe Gen5 et à une implémentation très bien conçue de la qualité de service du firmware, le DC3000ME offre l'une des bandes passantes et des IOPS par baie les plus élevées, tout en maintenant une faible latence. Ces deux éléments sont essentiels pour la fiabilité et la prévisibilité dans les environnements d'entreprise. Cela lui permet de répondre aux exigences des charges de travail des datacenters de pointe.

Louis Kaneshiro | Kingston Technology





Le rôle de Kingston dans l'optimisation des datacenters

Performances optimales

Les responsables de datacenters sont souvent confrontés à des problèmes de variabilité des performances, de complexité du déploiement et de gestion de la longévité des disques.

Le DC3000ME est fabriqué selon des spécifications de haute fiabilité afin d'offrir des performances constantes pour une disponibilité optimale. Il répond à ces défis grâce à une fonctionnalité de télémétrie initiée par l'hôte et le contrôleur afin de faciliter la surveillance en temps réel de l'état de santé et l'analyse prédictive des pannes.

À cela s'ajoute une conception robuste de qualité professionnelle, assortie d'une garantie de 5 ans et d'un MTBF de 2 millions d'heures afin d'éviter les temps d'arrêt et d'offrir des performances optimales.

Sécurité renforcée

Le DC3000ME est conforme à la norme TCG Opal 2.0 et inclut un chiffrement AES 256 bits pour une sécurité renforcée. Cette fonctionnalité garantit l'exclusion des utilisateurs non autorisés et empêche le vol des données du SSD si l'appareil est éjecté physiquement. Outre ses capacités de sécurité améliorées, le DC3000ME offre les avantages suivants :

- » **Protection matérielle contre les coupures d'alimentation :** l'intégrité des données est préservée même en cas de coupure d'alimentation imprévue.
- » **NVMe - Protection des données de bout en bout :** protection de bout en bout des données tout au long du parcours de stockage.
- » **Mise à jour du firmware sans réinitialisation :** facilite les mises à jour continues du firmware sans réinitialisation.

Économies à long terme

En tant qu'investissement, le DC3000ME est immédiatement rentable dans les déploiements périphériques de datacenters haute performance et les cas d'utilisation de formation IA avec contrôle QoS en temps réel. Tous deux sont essentiels pour la prise en charge des applications et l'évolutivité, ainsi que pour le traitement rapide des données.

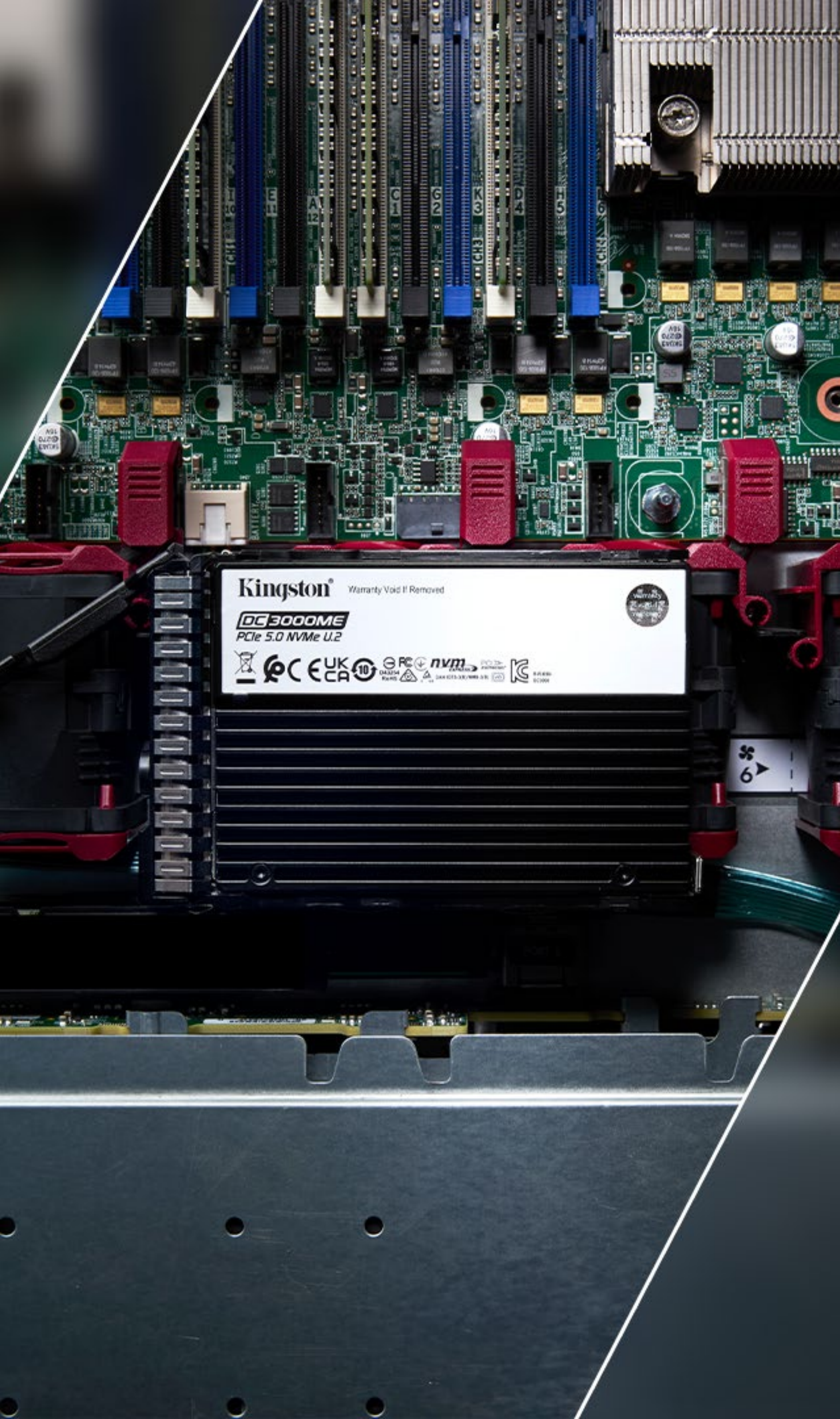
Malgré l'investissement initial élevé, les économies globales à long terme réalisées grâce à ses performances améliorées, sa consommation électrique réduite et ses temps d'arrêt moins fréquents en font un choix économique stratégique.



Le DC3000ME offre une valeur ajoutée significative sur le long terme grâce à sa faible consommation d'énergie, à la dégradation minimale de ses performances et à sa durée de vie prolongée.

Tony Hollingsbee | Kingston Technology Europe





Le rôle de Kingston dans l'optimisation des datacenters

Efficacité énergétique supérieure

La gestion de l'alimentation est une fonctionnalité unique du DC3000ME qui offre une grande flexibilité dans le contrôle des performances par watt, ce qui permet de réduire les coûts d'exploitation tout en diminuant l'empreinte carbone et en s'alignant sur les objectifs de développement durable.

La protection contre les coupures d'alimentation garantit la sauvegarde des données en cas de perte soudaine d'alimentation, même au niveau du SSD.



Parmi ses avantages stratégiques, citons la prise en charge du multitâche à haut débit sans interruption et une consommation d'énergie réduite pour des initiatives respectueuses de l'environnement.

Louis Kaneshiro | Kingston Technology



La série DC3000ME offre une efficacité énergétique supérieure, avec des performances de lecture séquentielle pouvant atteindre 970 Mo/s par watt. Grâce à une conception matérielle poussée et à l'optimisation du firmware, la série DC3000ME offre une meilleure utilisation du matériel et réduit au minimum son impact sur la dissipation thermique du serveur.

Louis Kaneshiro | Kingston Technology



Compatibilité étendue

La série DC3000ME est compatible avec différents serveurs et contrôleurs HBA OEM, prend en charge la norme PCIe Gen5 et est conforme à la spécification NVMe 2.0.

Elle prend en charge des fonctionnalités clés telles que NVMe-MI 2.0, un protocole standard qui permet la détection, la surveillance, la configuration et la mise à jour des périphériques NVMe dans différents environnements d'exploitation via la plateforme de gestion hors bande du serveur.

Par exemple, Dell iDRAC 9 et le dernier BMC Supermicro sont strictement conformes à la spécification NVMe MI 2.0, ce qui permet une intégration transparente avec les SSD NVMe compatibles MI tels que le DC3000ME.

Se préparer pour l'avenir

Pour pérenniser leur investissement, les entreprises doivent choisir des solutions SSD pour datacenters évolutives, capables de répondre aux exigences de performances actuelles et futures, tout en évitant la sous-utilisation ou la surutilisation des baies de lecteurs.

Kingston met son expertise à votre disposition pour vous aider à prendre les bonnes décisions. À mesure que l'IA, le ML et les nouvelles technologies progressent, nos solutions continuent de miser sur la croissance et l'efficacité de ces avancées prometteuses.

Ensemble, nous pouvons accélérer la transformation de votre datacenter grâce à notre expertise de leader du secteur et à nos solutions de stockage conçues pour s'adapter à l'évolution rapide des besoins.

Fondé sur l'engagement

Du Big Data aux appareils IoT, en passant par les ordinateurs portables, les PC et les technologies portables, Kingston Technology s'engage à fournir des solutions produits, des services et une assistance de premier ordre. Les principaux fabricants de PC et les fournisseurs mondiaux de services en ligne nous font confiance, et nous attachons une grande importance à nos partenariats à long terme qui nous aident à évoluer et à innover. Nous veillons à ce que chaque solution réponde aux normes les plus élevées en donnant la priorité à la qualité et au service client. À chaque étape, nous écoutons, apprenons et nous engageons avec nos clients et nos partenaires pour fournir des solutions qui ont un impact durable.

©2025 Kingston Technology Europe Co LLP et Kingston Digital Europe Co LLP, Kingston Court, Brooklands Close, Sunbury-on-Thames, Middlesex, TW16 7EP, Angleterre. Tél: +44 (0) 1932 738888 Fax: +44 (0) 1932 785469. Tous droits réservés. Toutes les marques commerciales et les marques déposées sont la propriété de leurs détenteurs respectifs



FONDÉ SUR L'ENGAGEMENT