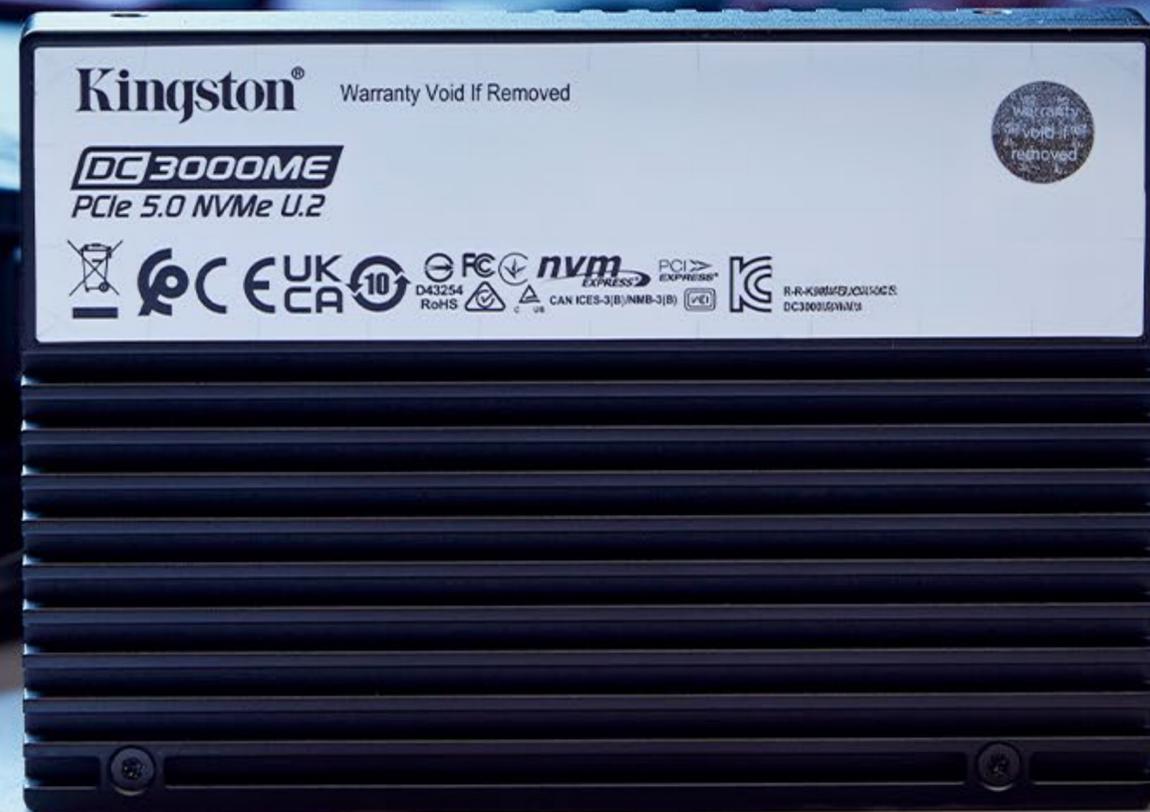


OTIMIZE O
ARMAZENAMENTO
DE DATA CENTER
COM SSDS
PCIE GEN5
PARA IA E
BIG DATA



Prefácio e índice

À medida que o mundo avança rumo a um futuro definido por dados, o foco se volta para a infraestrutura que os sustenta. Desde IA generativa até análises em tempo real, a demanda por velocidade, confiabilidade e eficiência está acelerando e, com ela, a necessidade de armazenamento em data centers de última geração. Como resultado, as Unidades de Estado Sólido (SSDs), especialmente as baseadas em PCIe Gen5, estão rapidamente se tornando a espinha dorsal dos ambientes de TI modernos.

Mas o que está impulsionando essa mudança? Como os gargalos de desempenho afetam a utilização da GPU? Por que o armazenamento corporativo energeticamente eficiente é tão crítico quanto os IOPS? E como os líderes de TI podem proteger a integridade dos dados enquanto preparam sua infraestrutura para o futuro?

Este eBook responde a essas perguntas e explora como melhorar o desempenho do armazenamento em data center, além de abordar as principais tendências do setor, desde o custo crescente da inatividade até a promessa de uma infraestrutura sustentável. Com insights de especialistas em armazenamento flash da Kingston, descubra como os SSDs avançados estão transformando os data centers para a era da IA e além.

Índice	Páginas
Colaboradores	3
O crescente fluxo de dados	4
Gargalos de desempenho e a necessidade de velocidade	5
O custo da inatividade e latência	6
Confiabilidade e integridade de dados	7
Gerenciando consumo de energia e sustentabilidade	8
Preparando sua infraestrutura de data center para o futuro	9
O papel da Kingston no fortalecimento dos data centers	10-12
Resumo e Sobre a Kingston	13



Colaboradores

Este eBook foi criado por dois especialistas da Kingston em tecnologias de armazenamento.



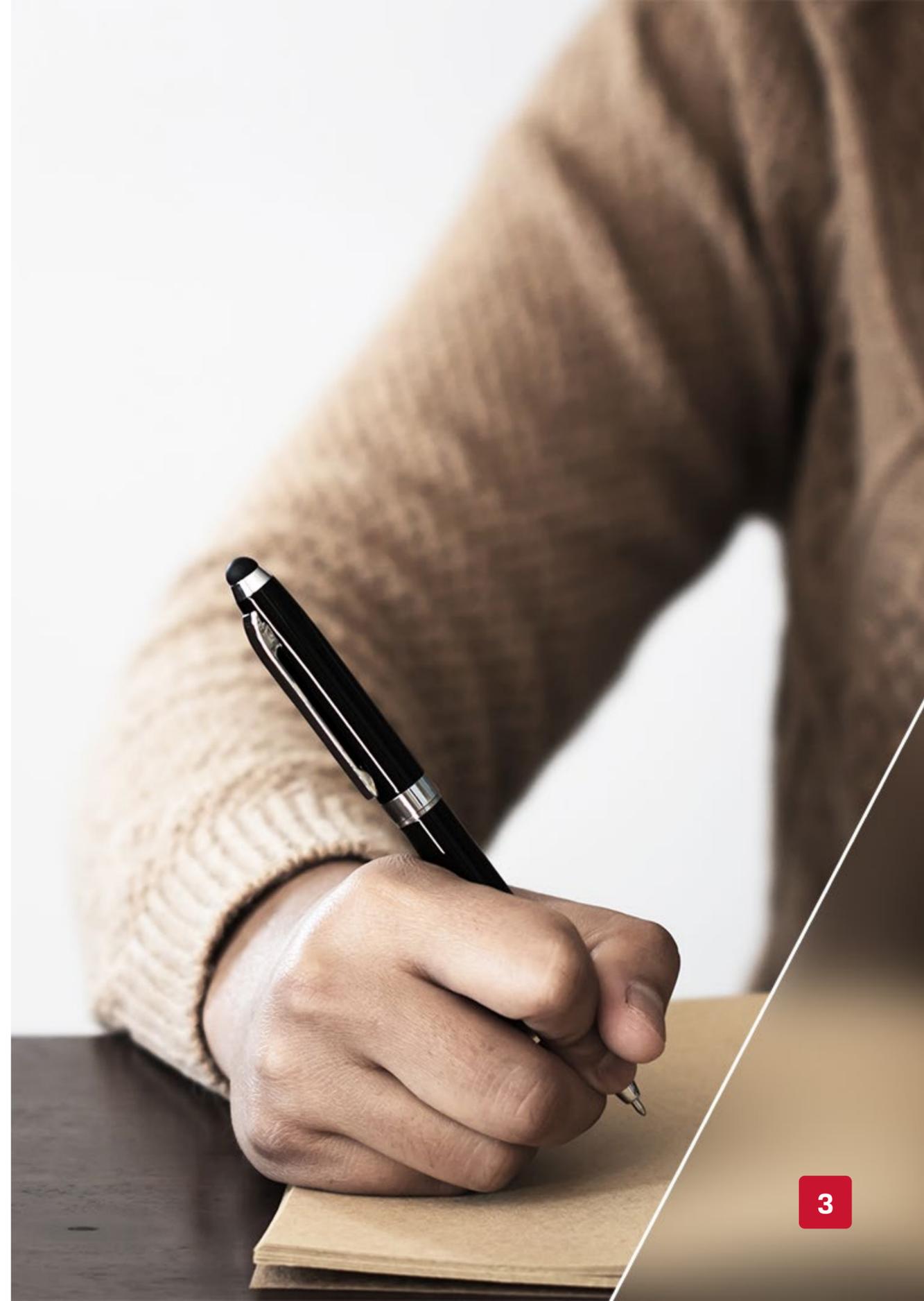
Louis Kaneshiro | Kingston Technology

Louis Kaneshiro é o Gerente de Tecnologia Sênior da Kingston Technology. Com 30 anos na Kingston, sendo os últimos 15 focados em SSD, ele liderou o Grupo de Recursos de Tecnologia, agora uma equipe global, antes de lançar o Departamento de Engenharia de Produto SSD (SPED). Antes de seu cargo atual, Louis atuou em suporte técnico e como FAE para a divisão OEM da Kingston.



Tony Hollingsbee | Kingston Technology Europe

Tony está na Kingston Technology há 23 anos, tendo ocupado vários cargos comerciais, inclusive como Gerente de Vendas Internas para as linhas de produtos de memória flash e SSD. Nos últimos 12 anos, atuou como Gerente de Negócios de SSD para a região EMEA, trabalhando em estreita colaboração com as equipes de vendas e marketing para apoiar o desenvolvimento de negócios e impulsionar o crescimento regional.



O crescente fluxo de dados

A quantidade exponencial de dados criados e consumidos globalmente deve ultrapassar 394 zettabytes até 2028¹. Grande parte dessa explosão de dados é impulsionada pelos avanços em inteligência artificial (IA), Internet das Coisas (IoT) e consumo de mídia digital. Aplicações de IA e machine learning geram e consomem conjuntos de dados massivos. Enquanto tecnologias emergentes, como veículos autônomos e cidades inteligentes, impulsionam a geração de dados em tempo real em uma escala sem precedentes.

Essas tendências exigem sistemas de armazenamento de alta taxa de transferência e baixa latência. Cargas de trabalho corporativas, aplicações nativas da nuvem e conformidade regulatória também contribuem para a crescente demanda por armazenamento escalável e seguro. Para acomodar esse crescimento, os data centers estão evoluindo suas infraestruturas de armazenamento. O surgimento do Enterprise and Data Center Standard Form Factor (EDSFF) está ganhando força. E tecnologias inovadoras de armazenamento, como soluções de arquivamento baseadas em cerâmica com capacidade de retenção de dados superior a 5.000 anos, estão sendo desenvolvidas para atender às demandas de preservação de dados em longo prazo.

Embora o cenário atual de formatos de armazenamento em data centers seja amplo e variado, fatores tradicionais como os SSDs U.2, conhecidos por sua interface PCIe e adequação em ambientes de alto desempenho, foram amplamente adotados. Isso é especialmente verdadeiro entre OEMs de servidores e fabricantes de chassi. Esses avanços visam, coletivamente, suprir as crescentes demandas de armazenamento dos data centers modernos diante da proliferação implacável de dados.



SSDs como o DC3000ME são fundamentais para lidar com tendências computacionais emergentes, especialmente em IA, aprendizado de máquina e computação de borda. Nosso projeto antecipa a necessidade de soluções de armazenamento com maior largura de banda, menor latência e maior eficiência energética.

Tony Hollingsbee | Kingston Technology Europe



Embora os SSDs E1.x e E3.x estejam ganhando apoio, o formato U.2 é, de longe, o dominante, com mais de 60% dos petabytes em servidores residindo em SSDs U.2.

Louis Kaneshiro | Kingston Technology



1. Volume de dados/informações criados, capturados, copiados e consumidos globalmente de 2010 a 2023, com projeções de 2024 a 2028 (em zettabytes). <https://www.statista.com/statistics/871513/worldwide-data-created/>



“

Com a IA em destaque agora, a necessidade de armazenamento de maior desempenho para maximizar a utilização da GPU é mais importante do que nunca.

Louis Kaneshiro | Kingston Technology

”

Gargalos de desempenho e a necessidade de velocidade

À medida que os volumes de dados aumentam e as aplicações se tornam mais complexas, os gestores de data centers enfrentam desafios crescentes. Um dos problemas mais persistentes é a variabilidade de desempenho. Muitas soluções de armazenamento em data center corporativo não conseguem fornecer taxa de transferência e latência consistentes, levando a ineficiências no desempenho das aplicações e na utilização de recursos.

Implantações complexas complicam ainda mais esse cenário, muitas vezes exigindo configurações personalizadas, ajustes de firmware e orquestração intrincada para atender às necessidades específicas das cargas de trabalho. Além disso, gerenciar a longevidade dos drives é uma preocupação crescente. Principalmente porque cargas de trabalho com alta gravação, como registro em log, cache e treinamento de IA, exercem enorme pressão sobre os SSDs, acelerando o desgaste e aumentando o risco de inatividade.

As cargas de trabalho de IA dependem fortemente de GPUs para computação, mas, a menos que o armazenamento subjacente consiga acompanhar, a utilização da GPU despenca, resultando em investimentos desperdiçados e metas de desempenho não alcançadas. SSDs de alto desempenho para cargas de trabalho de IA, com latência ultrabaixa e alto rendimento, são essenciais para sustentar os pipelines de dados que alimentam as GPUs em tempo real.

Somando-se a esses desafios, a necessidade de velocidade nos data centers nunca foi tão grande. Cargas de trabalho modernas, especialmente IA, aprendizado de máquina e análises em tempo real, exigem acesso instantâneo a grandes volumes de dados. Qualquer latência na camada de armazenamento pode interromper o processamento, atrasar insights e reduzir a eficiência geral do sistema. Armazenamento de alta velocidade não é mais um luxo; é uma necessidade para recursos de computação de alto desempenho, como GPUs e TPUs.

0 custo da inatividade e latência

A inatividade e a latência são preocupações adicionais para os líderes de TI, com potencial para perdas financeiras substanciais e interrupções operacionais. Estudos recentes relatam que o custo da inatividade por hora ultrapassa US\$ 300.000 para 90% das empresas, com 41% dizendo que a inatividade por hora custa de US\$ 1 milhão até mais de US\$ 5 milhões².

A latência agrava ainda mais esses desafios, prejudicando a velocidade de acesso aos dados, o que é particularmente prejudicial em ambientes intensivos em dados. Atrasos na recuperação de dados degradam a experiência do usuário, reduzem a produtividade e levam a perdas de receita. No trading de alta frequência, por exemplo, a latência é um fator crítico, onde até mesmo alguns milissegundos podem determinar milhões em lucro ou perda.

SSDs avançados oferecem uma solução atraente para esses problemas. Aproveitando tecnologias como interfaces PCIe Gen5 e arquiteturas avançadas 3D NAND, os SSDs modernos oferecem velocidades de leitura e gravação significativamente mais rápidas em comparação com os discos rígidos tradicionais (HDDs) e SSDs SATA.

Esse desempenho aprimorado não apenas acelera o acesso aos dados, mas também melhora a confiabilidade do sistema, reduzindo a probabilidade e o impacto da inatividade. Ao integrar esses SSDs avançados em sua infraestrutura, as empresas podem alcançar maior eficiência operacional e resiliência.



Os SSDs de última geração para computação de borda oferecem processamento de dados de baixa latência e alta velocidade, suportando operações em tempo real críticas para sistemas distribuídos.

Louis Kaneshiro | Kingston Technology



2. O custo da inatividade por hora ultrapassa US\$ 300.000 para 90% das empresas, 41% das empresas dizem que a inatividade por hora custa de US\$ 1 milhão até mais de US\$ 5 milhões. <https://itic-corp.com/itic-2024-hourly-cost-of-downtime-report/>

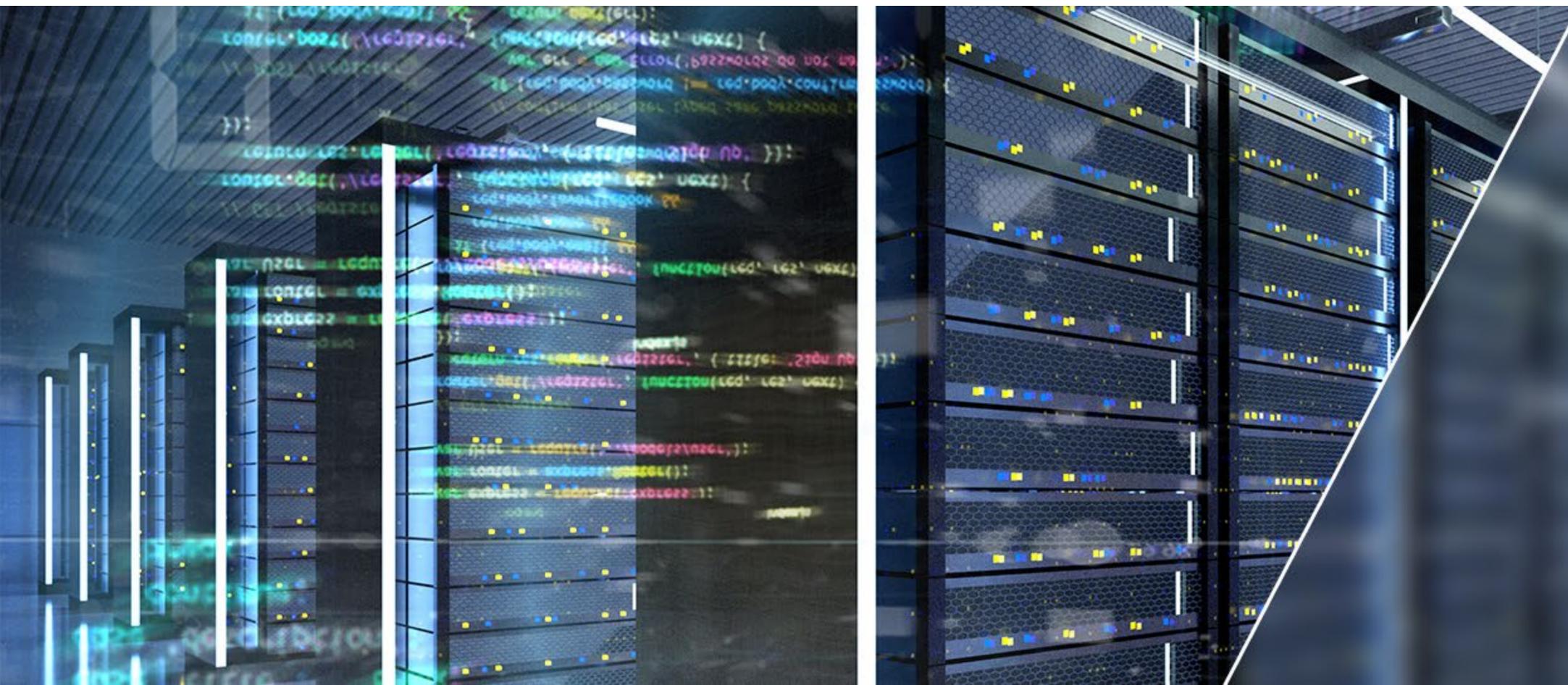


Confiabilidade e integridade dos dados

Na economia digital atual, confiabilidade e integridade de dados são inegociáveis para data centers que gerenciam cargas de trabalho críticas. À medida que as empresas dependem cada vez mais de análises em tempo real, aplicações nativas em nuvem e sistemas impulsionados por IA, até mesmo pequenas perdas ou corrupções de dados podem levar a contratempos operacionais e danos à reputação.

A confiabilidade dos dados depende fortemente da robustez da infraestrutura de armazenamento. Fatores como resistência do drive, mecanismos de correção de erros e capacidades de failover são vitais para manter a disponibilidade consistente e a saúde dos dados. A mídia de armazenamento não deve apenas resistir ao desgaste físico, mas também detectar e corrigir erros em nível de bit antes que afetem as aplicações.

Os SSDs corporativos, com tecnologias como proteção de dados ponta a ponta, proteção contra falha de energia e correção avançada de erros (como LDPC), são projetados para proteger os dados mesmo sob cargas de trabalho extremas. E com a crescente prevalência de ransomware e ameaças cibernéticas, preservar a integridade dos dados durante ataques é igualmente fundamental. Os SSDs de última geração – desenvolvidos para resistência, consistência e resiliência – fornecem a camada fundamental de confiança que os data centers modernos precisam. Ao priorizar confiabilidade e integridade de dados, as organizações podem proteger seu ativo mais valioso: seus dados.



“

Com demandas de computação crescentes, a escalabilidade dos SSDs Gen5 permite que os data centers se adaptem e atendam às necessidades do amanhã. Recursos como proteção contra falha de energia assistida por hardware garantem a integridade dos dados; essencial em aplicações críticas, onde a preservação de dados é prioridade.

Louis Kaneshiro | Kingston Technology

”



Gerenciando consumo de energia e sustentabilidade

Juntamente com a garantia de confiabilidade e integridade de dados, reduzir o consumo de energia dos data centers é uma das principais prioridades dos líderes de TI. À medida que metas de sustentabilidade e eficiência operacional convergem, e os volumes globais de dados disparam, também aumenta a energia necessária para processar, armazenar e mover esses dados.

Em 2023, a demanda global por data centers ficou em torno de 340 TWh, representando aproximadamente 1,3% do uso mundial de eletricidade. Múltiplas fontes preveem que a demanda aumentará drasticamente até 2030, ampliando a pressão sobre a infraestrutura de energia e os esforços de sustentabilidade³.

Reduzir o consumo de energia não se trata apenas de cumprir regulamentações ambientais e metas corporativas de ESG (Meio Ambiente, Social e Governo). É fundamental para reduzir custos com utilitários e garantir a viabilidade de um data center. Como resultado, a otimização de armazenamento é uma área-chave de foco para os gerentes de data center atuais.

Os discos rígidos tradicionais (HDDs) consomem mais energia e geram mais calor em comparação com SSDs SATA. Em contraste, embora os SSDs PCIe NVMe de próxima geração estejam mais próximos dos HDDs em termos de consumo de energia e geração de calor, eles oferecem uma eficiência energética significativamente melhor quando medidos em desempenho por watt. Graças ao impressionante número de IOPS que os SSDs PCIe NVMe podem suportar, os dados podem ser acessados e processados rapidamente pelos sistemas hospedeiros, reduzindo os tempos de operação e impactando positivamente o uso geral de energia.

SSDs PCIe NVMe de alta capacidade também permitem armazenamento de maior densidade em espaços menores. Isso não apenas apoia operações mais sustentáveis, mas também melhora a escalabilidade da infraestrutura de dados. Dessa forma, os SSDs avançados são fundamentais para um projeto de data center sustentável e de alto desempenho.

A eficiência energética dos SSDs é especialmente útil, pois mantém alto desempenho enquanto reduzem o consumo de energia e a pegada de carbono geral – uma consideração fundamental para dispositivos que provavelmente operarão em implantações de data center de médio ou alto nível.

Louis Kaneshiro | Kingston Technology

3. O crescimento na demanda global por energia disparou em 2024, atingindo quase o dobro da média recente. <https://www.iea.org/news/growth-in-global-energy-demand-surged-in-2024-to-almost-twice-its-recent-average>

Preparando sua infraestrutura de data center para o futuro

À medida que as cargas de trabalho se tornam mais complexas e os volumes de dados aumentam, a adoção de tecnologias escaláveis e de alto desempenho torna-se fundamental. Os SSDs PCIe 5.0 são um divisor de águas nesse espaço, oferecendo o dobro da largura de banda em relação à Gen 4 e reduzindo significativamente a latência. Eles permitem acesso mais rápido aos dados, melhoram a utilização da GPU para cargas de trabalho de IA e suportam implantações de armazenamento mais densas.

A integração de SSDs e arquiteturas de servidor compatíveis com a Gen 5 garante que a infraestrutura possa lidar com demandas futuras, seja em análises em tempo real, aprendizado de máquina ou computação de borda, sem exigir constantes e custosas reformas de hardware.

Ao considerar SSDs Gen 5 para data centers, as empresas devem começar avaliando seus aplicativos e necessidades de armazenamento para determinar se esse nível de desempenho é necessário. Esse processo deve começar com a avaliação do arquiteto de armazenamento sobre os requisitos de desempenho em um, dois ou cinco anos.

Diferentes ambientes de data center e previsões de expansão (scale-up/scale-out) para crescimento de armazenamento devem ser considerados, juntamente com medição e avaliação de ROI. Para isso, a telemetria de aplicativos pode ajudar a coletar métricas de desempenho detalhadas e expectativas de crescimento de capacidade de armazenamento.

A infraestrutura de hardware atual deve ser levada em conta, assim como a compatibilidade com infraestruturas PCIe 5.0 existentes, para entender os requisitos de fornecimento de energia e validar soluções de gerenciamento térmico.



O custo total de propriedade deve ser avaliado não apenas pelo preço inicial de compra, mas pelo desempenho por watt, métricas de confiabilidade e eficiência operacional total.

Tony Hollingsbee | Kingston Technology Europe



Ao examinar esses fatores em profundidade e considerar estratégias de preparação para o futuro/evitar armadilhas, as organizações podem tomar decisões informadas que garantam estar adequadamente equipadas para enfrentar os desafios atuais e os avanços tecnológicos de amanhã.

Louis Kaneshiro | Kingston Technology





O papel da Kingston no fortalecimento dos data centers

SSDs como o [DC3000ME da Kingston](#) estão no coração do aprimoramento do desempenho de servidores e GPUs. Eles oferecem a capacidade de reduzir tempo de inatividade e lidar com computação de borda, infraestrutura de IA e processos computacionais intensos com facilidade.

Representando um grande salto no desempenho de armazenamento em data centers, o DC3000ME foi projetado especificamente para atender às demandas exponenciais de cargas de trabalho de IA de próxima geração e computação de alto desempenho.

Nossas principais inovações de engenharia incluem uma interface PCIe 5.0 revolucionária com tecnologia de controlador avançada. Este recurso oferece velocidades de leitura/gravação sem precedentes, gerenciamento térmico integrado e um formato otimizado para ambientes de servidor de alta densidade. Tais avanços destacam nosso compromisso em moldar um ecossistema tecnológico robusto e ágil para beneficiar as futuras gerações de data centers. É assim:

Desenvolvido para cargas de trabalho pesadas

Para cargas de trabalho de IA e ML, o DC3000ME com sua interface Gen5 oferece rendimento de leitura superior e IOPS de leitura. Isso significa que mesmo um único drive DC3000ME pode entregar IO de forma eficiente para múltiplas GPUs, saturando totalmente o desempenho. O resultado são tempos de treinamento mais rápidos e menos SSDs necessários, especialmente em comparação com a Gen4, para acompanhar múltiplas GPUs, reduzindo custos de infraestrutura sem comprometer o desempenho.

O DC3000ME complementa a infraestrutura de IA com sua alta largura de banda, lidando facilmente com cargas de trabalho intensivas em dados e permitindo utilização contínua das GPUs. A capacidade de atender múltiplas GPUs simultaneamente

acelera atividades de treinamento e inferência de IA, tornando-as mais eficientes e produtivas. Com demandas de computação crescentes, a escalabilidade dos SSDs Gen5 permite que os data centers se adaptem e atendam às necessidades do amanhã.

“

O DC3000ME apresenta métricas de desempenho excepcionais, com velocidades de leitura sequencial de até 14.000 MB/s e gravação de até 10.000 MB/s. Ele oferece diferentes estágios de energia, permitindo flexibilidade no consumo de energia com desempenho escalável.

Tony Hollingsbee | Kingston Technology Europe

”

“

Com a implementação da mais recente tecnologia PCIe NVMe Gen5 e um firmware de Qualidade de Serviço (QoS) extremamente bem-otimizado, o DC3000ME oferece uma das maiores larguras de banda e IOPS por compartimento, mantendo baixa latência de forma sustentada. Ambos são pilares essenciais para confiabilidade e previsibilidade em ambientes corporativos. Isso garante que ele possa acompanhar as demandas das cargas de trabalho mais avançadas em data centers.

Louis Kaneshiro | Kingston Technology

”



O papel da Kingston no fortalecimento dos data centers

Desempenho otimizado

Gerentes de data center frequentemente enfrentam variabilidade de desempenho, implantação complexa e gestão da longevidade dos drives.

O DC3000ME é fabricado com especificações de alta confiabilidade para oferecer desempenho consistente e tempo de atividade ideal. Ele aborda esses desafios com funcionalidades de telemetria iniciadas pelo host e pelo controlador, facilitando monitoramento em tempo real e análise preditiva de falhas.

Isso é combinado com um design robusto de nível empresarial, garantia de 5 anos e MTBF de 2 milhões de horas para evitar inatividade e entregar desempenho otimizado.

Maior segurança

O DC3000ME é compatível com TCG Opal 2.0 e inclui criptografia AES de 256 bits para segurança robusta. Esse recurso garante que usuários não autorizados sejam bloqueados e que os dados do drive não possam ser roubados se o dispositivo for fisicamente removido. Além de suas capacidades de segurança aprimoradas, o DC3000ME oferece:

- » **Proteção contra falhas de energia baseada em hardware:** A integridade dos dados é mantida mesmo em caso de falha de energia inesperada.
- » **Proteção de dados de ponta a ponta NVMe:** Fornece proteção integral dos dados ao longo de todo o caminho de armazenamento.
- » **Atualização de firmware sem reinicialização:** Facilita atualizações contínuas de firmware sem necessidade de reinicialização.

Economia de custo a longo prazo

Como investimento, o DC3000ME compensa imediatamente em implantações de borda de data center de alto desempenho e casos de uso de treinamento de IA com controle de QoS em tempo real. Ambos são essenciais para suporte a aplicativos e escalabilidade, além de processamento de dados em alta velocidade.

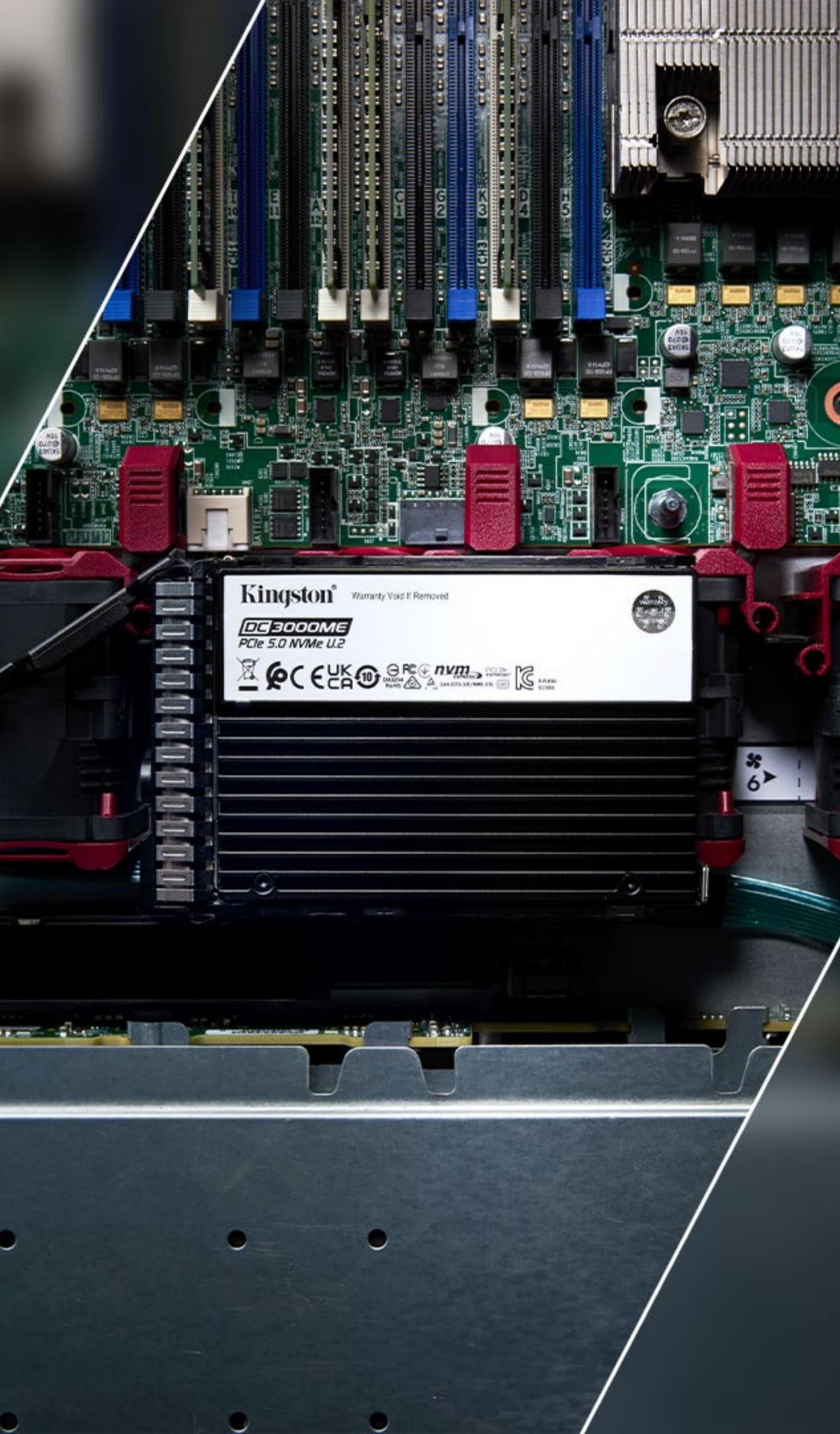
Apesar do alto investimento inicial, economias significativas a longo prazo podem ser alcançadas devido ao seu desempenho aprimorado, requisitos de energia reduzidos e menos tempo de inatividade, tornando-o uma escolha econômica estratégica.



O DC3000ME oferece valor significativo a longo prazo por meio de menor consumo de energia, degradação mínima de desempenho e vida operacional estendida.

Tony Hollingsbee | Kingston Technology Europe





O papel da Kingston no fortalecimento dos data centers

Eficiência energética superior

O gerenciamento de energia é um recurso exclusivo do DC3000ME, oferecendo flexibilidade no controle de desempenho por watt de utilização do drive, reduzindo custos operacionais e pegada de carbono, alinhando-se com metas sustentáveis.

A proteção contra falhas de energia garante que os dados sejam protegidos em caso de falha repentina de energia, mesmo no nível do drive.



Vantagens estratégicas incluem suporte de multitarefa de alta largura de banda sem falhas e consumo de energia reduzido para iniciativas ambientalmente sustentáveis.

Louis Kaneshiro | Kingston Technology



A série DC3000ME apresenta eficiência energética superior, oferecendo até 970MB/s de desempenho de leitura sequencial por watt. Por meio de design de hardware avançado e otimização de firmware, a série DC3000ME alcança maior utilização de hardware e minimiza seu impacto na dissipação de calor do servidor.

Louis Kaneshiro | Kingston Technology



Ampla compatibilidade

A série DC3000ME é compatível com diferentes servidores e HBAs OEM, suportando o padrão PCIe Gen5 e em conformidade com a especificação NVMe 2.0.

Ela suporta recursos essenciais como NVMe-MI 2.0, um protocolo padrão do setor que permite descoberta, monitoramento, configuração e atualização de dispositivos NVMe em diferentes ambientes operacionais pela plataforma de gerenciamento out-of-band do servidor.

Por exemplo, o Dell iDRAC 9 e o mais recente Supermicro BMC estão estritamente em conformidade com a especificação NVMe MI 2.0, permitindo integração perfeita com drives NVMe compatíveis com MI, como o DC3000ME.

Construindo para o futuro

Para proteger seu investimento, as organizações devem escolher soluções de SSD escaláveis para data centers que atendam às demandas de desempenho atuais e futuras – evitando subutilização ou superutilização de compartimentos de drive.

A Kingston oferece a expertise necessária para orientar suas decisões. À medida que IA, ML e novas tecnologias avançam, nossas soluções permanecem focadas em apoiar o crescimento e a eficácia desses avanços poderosos.

Juntos, podemos acelerar a transformação do seu data center com a expertise de liderança no setor e soluções de armazenamento projetadas para acompanhar mudanças rápidas.

Built on Commitment

De grandes volumes de dados a dispositivos de Internet das Coisas, incluindo laptops, PCs e tecnologia wearable, a Kingston Technology dedica-se a fornecer soluções de produtos de primeira linha, serviços e suporte. Com a confiança dos principais fabricantes de PCs e fornecedores globais de nuvem, valorizamos as nossas parcerias de longo prazo que nos ajudam a evoluir e inovar. Garantimos que cada solução atenda aos mais altos padrões, priorizando qualidade e atendimento ao cliente. A cada passo, ouvimos, aprendemos e nos envolvemos com os nossos clientes e parceiros para oferecer soluções que tenham um impacto duradouro.

©2025 Kingston Technology Corporation, 17600 Newhope Street, Fountain Valley, CA 92708 USA.

Todos os direitos reservados. Todas as marcas ou marcas registradas pertencem a seus respectivos proprietários.



 **Kingston**[®]
TECHNOLOGY
BUILT ON COMMITMENT

BUILT ON COMMITMENT