Comunicado de Imprensa

**Microsoft anuncia novo sistema de arrefecimento microfluídico de chips de IA três vezes mais eficaz**

**Lisboa, 01 de outubro de 2025** – Os chips utilizados em centros de dados para executar os mais recentes avanços em inteligência artificial (IA) geram muito mais calor do que as gerações anteriores de silício, e quem já teve o telemóvel ou o portátil a sobreaquecer sabe que os componentes eletrónicos não funcionam bem com temperaturas elevadas. Perante a crescente procura por IA e os novos designs de chips, a tecnologia de arrefecimento atual poderá limitar o progresso dentro de poucos anos.

Para ajudar a resolver este problema, a Microsoft testou, com sucesso, um novo sistema de arrefecimento que remove o calor até três vezes melhor do que as placas frias, uma tecnologia avançada de arrefecimento amplamente utilizada atualmente. Este sistema recorre à microfluídica, uma técnica que conduz o líquido refrigerante diretamente para o interior do silício, onde o calor é gerado. Canais microscópicos são gravados na parte traseira do chip de silício, criando sulcos que permitem ao líquido refrigerante fluir diretamente sobre o chip e remover o calor de forma mais eficiente. A equipa também utilizou IA para identificar as assinaturas térmicas únicas de cada chip e direcionar o refrigerante com maior precisão.

Os investigadores afirmam que a microfluídica poderá aumentar a eficiência e melhorar a sustentabilidade dos chips de IA da próxima geração. Atualmente, a maioria das GPU em funcionamento nos centros de dados são arrefecidas com placas frias, que estão separadas da fonte de calor por várias camadas, limitando a quantidade de calor que conseguem dissipar.

À medida que cada nova geração de chips de IA se torna mais potente, gera também mais calor. Dentro de cinco anos, “se continuarmos a depender fortemente da tecnologia tradicional de placas frias, ficamos bloqueados”, afirma Sashi Majety, gestora sénior de programas técnicos na área de Operações e Inovação na Cloud da Microsoft.

Neste contexto, a Microsoft conseguiu desenvolver um sistema de arrefecimento microfluídico integrado no chip, capaz de arrefecer eficazmente um servidor a executar serviços nucleares durante uma reunião simulada no Microsoft Teams. “A microfluídica permitirá designs mais densos em termos de potência, que viabilizam mais funcionalidades valorizadas pelos clientes e oferecem melhor desempenho em menos espaço”, refere Judy Priest, vice-presidente corporativa e diretora técnica de Operações e Inovação na Cloud da Microsoft. “Mas era necessário provar que a tecnologia e o design funcionavam, e o passo seguinte foi testar a fiabilidade”, acrescenta.

Os testes em laboratório revelaram que a microfluídica teve um desempenho até três vezes superior ao das placas frias na remoção de calor, dependendo das cargas de trabalho e das configurações envolvidas. A microfluídica também reduziu em 65% o aumento máximo de temperatura do silício dentro da GPU, embora este valor varie consoante o tipo de chip. A equipa acredita que esta tecnologia de arrefecimento avançada melhorará também a eficácia do uso de energia (PUE), um indicador-chave da eficiência energética de um centro de dados, e reduzirá os custos operacionais.

**Utilizar IA para imitar a natureza**

A microfluídica não é um conceito novo, mas torná-la funcional tem sido um desafio para o setor. “O pensamento sistémico é crucial no desenvolvimento de tecnologias como a microfluídica. É necessário compreender as interações entre sistemas — silício, refrigerante, servidor e centro de dados — para tirar o máximo partido da tecnologia”, explica Husam Alissa, diretor de tecnologia de sistemas na área de Operações e Inovação na Cloud da Microsoft.

Acertar nos sulcos já é difícil, uma vez que as dimensões dos microcanais são semelhantes às de um fio de cabelo humano, o que significa que não há margem para erro. Durante a prototipagem, a Microsoft colaborou com a startup suíça [Corintis](https://corintis.com/) para utilizar IA na otimização de um design bioinspirado que arrefece os pontos quentes dos chips de forma mais eficiente do que os canais verticais, também testados. Este design assemelha-se às veias de uma folha ou às asas de uma borboleta, visto que a natureza é perita em encontrar os caminhos mais eficientes para distribuir o que é necessário.

A microfluídica exige mais do que um design inovador de canais. É um desafio de engenharia complexo, pelo que foi necessário garantir que os canais fossem suficientemente profundos para permitir a circulação adequada do líquido refrigerante sem obstruções, mas não tão profundos que enfraquecessem o silício ao ponto de se partir, tendo a equipa produzido quatro iterações de design no último ano.

A microfluídica exigiu também o desenvolvimento de um encapsulamento estanque para o chip, a formulação do melhor refrigerante, a experimentação de diferentes métodos de gravação e a criação de um processo passo a passo para integrar a gravação na produção dos chips.

Este avanço é um dos exemplos de como a Microsoft está a investir e a inovar na infraestrutura para responder à procura por serviços e capacidades de IA, sendo que a empresa planeia alocar mais de 30 mil milhões de dólares em investimento de capital neste trimestre.

Estes investimentos incluem o desenvolvimento da [família de chips Cobalt e Maia](https://news.microsoft.com/source/features/ai/in-house-chips-silicon-to-service-to-meet-ai-demand/?msockid=0d710b8d313360371e1f1f27301e6148), concebidos especificamente para executar cargas de trabalho da Microsoft e dos seus clientes de forma mais eficiente. Desde que a Microsoft implementou o chip Cobalt 100[, tanto a empresa como os seus clientes têm beneficiado](https://aka.ms/cobalt100customers) da sua capacidade de computação eficiente em termos energéticos, escalabilidade e desempenho.

Os chips são apenas uma peça do puzzle, já que o silício funciona dentro de um sistema complexo de placas, *racks* e servidores num centro de dados. A abordagem sistémica da Microsoft implica afinar cada componente para que todos funcionem em conjunto e maximizem o desempenho e a eficiência, e uma parte fundamental desse esforço é o desenvolvimento de técnicas de arrefecimento de nova geração, como a microfluídica.

Como próximo passo, a Microsoft continua a investigar como integrar o arrefecimento microfluídico nas futuras gerações dos seus chips proprietários e continuará também a colaborar com parceiros de fabrico e silício para levar a microfluídica à produção em larga escala nos seus centros de dados.

“O hardware é a base dos nossos serviços”, afirma Jim Kleewein, técnico especializado na área de Gestão Nuclear do Microsoft 365. “Todos temos interesse nessa base, na sua fiabilidade, custo, velocidade, consistência de comportamento e sustentabilidade, só para citar alguns. A microfluídica melhora todos esses aspetos: custo, fiabilidade, velocidade, consistência e sustentabilidade.”

**Vantagens da microfluídica**

Uma simples chamada no Microsoft Teams ilustra as vantagens que o arrefecimento microfluídico pode oferecer. O Teams não é um serviço único, mas um aglomerado de cerca de 300 serviços diferentes que funcionam em conjunto. Um conecta o utilizador à reunião, outro aloja a reunião, outro guarda o chat, outro funde os fluxos de áudio para que todos sejam ouvidos, outro grava, outro transcreve.

“Cada serviço tem características diferentes e sobrecarrega partes distintas do servidor”, explica Kleewein. “Quanto mais utilizado for um servidor, mais calor gera, o que faz sentido.”

Por exemplo, a maioria das chamadas no Teams começa à hora certa ou à meia hora. O controlador de chamadas fica muito ocupado cerca de cinco minutos antes e três minutos depois desses momentos, e pouco ocupado no restante tempo. Existem duas formas de lidar com picos de procura: instalar muita capacidade extra dispendiosa que não é usada na maior parte do tempo ou forçar os servidores a trabalhar mais intensamente, o que se chama *overclocking.* Como o *overclocking* aquece ainda mais os chips, não pode ser feito em excesso sob pena de os danificar.

“Sempre que temos cargas de trabalho irregulares, queremos poder fazer *overclock*. A microfluídica permitir-nos-ia fazêe-lo sem receio de derreter o chip, porque é um sistema de arrefecimento mais eficiente”, afirma Kleewein. “Há vantagens em termos de custo e fiabilidade. E de velocidade, porque podemos fazer *overclock*.”

**O arrefecimento como parte de uma visão mais ampla**

A microfluídica faz parte de uma iniciativa mais abrangente da Microsoft para desenvolver técnicas de arrefecimento avançadas e otimizar cada componente da *stack cloud.* Tradicionalmente, os centros de dados têm sido arrefecidos com ar impulsionado por grandes ventoinhas, mas os líquidos conduzem o calor de forma muito mais eficiente do que o ar.

Uma das [formas de arrefecimento líquido que a Microsoft já implementou](https://news.microsoft.com/source/features/ai/in-house-chips-silicon-to-service-to-meet-ai-demand/?msockid=0d710b8d313360371e1f1f27301e6148) nos seus centros de dados são as placas frias. Estas placas assentam sobre os chips, com líquido frio a circular através de canais internos nas placas para absorver o calor dos chips por baixo, sendo depois expelido como líquido quente para ser novamente arrefecido.

Os chips são encapsulados com várias camadas de materiais que ajudam a dissipar o calor dos pontos mais quentes e a protegê-los. No entanto, esses materiais funcionam também como isolantes, limitando o desempenho das placas frias ao reterem o calor e impedirem a entrada do frio. Prevê-se que as futuras gerações de chips otimizados para IA sejam ainda mais potentes e que atinjam temperaturas demasiado elevadas para serem arrefecidas apenas com placas frias.

Arrefecer os chips diretamente através de canais microfluídicos é muito mais eficiente não só na remoção do calor, mas também no funcionamento global do sistema. Com todas essas camadas de isolamento removidas e o refrigerante a tocar diretamente no silício quente, o líquido refrigerante já não precisa de estar tão frio para cumprir a sua função. Isso poupa energia que, de outra forma, seria necessária para arrefecer o líquido, e oferece um desempenho superior ao das placas frias. A tecnologia de microfluídica permite também uma utilização mais eficiente do calor residual.

A Microsoft pretende ainda otimizar as operações dos centros de dados através de software e outras abordagens. “Se o arrefecimento microfluídico conseguir usar menos energia para arrefecer os centros de dados, isso reduzirá a pressão sobre as redes elétricas das comunidades vizinhas”, afirmou Ricardo Bianchini, técnico especializado da Microsoft e vice-presidente corporativo da Azure, especializado em eficiência computacional.

O calor também impõe limites ao design dos centros de dados. Uma das vantagens de um centro de dados é que os servidores estão fisicamente próximos uns dos outros. A distância atrasa a comunicação entre servidores, algo que se designa por latência. Mas os servidores atuais só podem ser agrupados até certo ponto antes de o calor se tornar um problema. A microfluídica permitiria aumentar a densidade de servidores nos centros de dados, o que significa que estes poderiam aumentar a capacidade de computação sem necessidade de construir novos edifícios.

**O futuro da inovação em chips**

A microfluídica também tem potencial para abrir caminho a arquiteturas de chips completamente novas, como os chips 3D. Tal como colocar servidores próximos reduz a latência, empilhar chips reduz ainda mais. Este tipo de arquitetura 3D é difícil de construir devido ao calor que gera.

No entanto, a microfluídica leva o refrigerante extremamente perto do ponto de consumo de energia, pelo que “podemos fazer o líquido circular através do chip”, como seria o caso dos designs 3D, explicou Bianchini, o que exigiria um design microfluídico diferente, com pinos cilíndricos entre os chips empilhados, como pilares num parque de estacionamento de vários andares, com o líquido a circular à volta deles.

“Sempre que conseguimos fazer as coisas de forma mais eficiente e simplificada, isso abre oportunidades para novas inovações, onde podemos explorar novas arquiteturas de chips”, afirma Priest.

Eliminar o limite imposto pelo calor poderá também permitir mais chips num *rack* de centro de dados ou mais núcleos num chip, o que aumentaria a velocidade e permitiria centros de dados mais pequenos, mas mais potentes.

Ao demonstrar como técnicas de arrefecimento como a microfluídica podem ser aplicadas com sucesso, a Microsoft espera ajudar a abrir caminho para chips de próxima geração mais eficientes e sustentáveis em todo o setor.

“Queremos que a microfluídica se torne algo que todos façam, não apenas algo que nós fazemos”, conclui Kleewein. “Quanto mais pessoas a adotarem, mais depressa a tecnologia se desenvolverá, e melhor será para nós, para os nossos clientes, para todos.”

Saiba mais [aqui](https://datacenters.microsoft.com/wp-content/uploads/2025/09/Microfluidic-Cooling-Infographic.pdf).

**CONTACTOS**

|  |  |  |
| --- | --- | --- |
| **Catarina Brito** | **Inês Filipe** |  |
| [catarina.brito@lift.com.pt](mailto:catarina.brito@lift.com.pt) | [ines.filipe@lift.com.pt](mailto:ines.filipe@lift.com.pt) |  |
| 914 310 661 | 910 283 054 |  |

**SOBRE A MICROSOFT**

Microsoft (Nasdaq "MSFT" @microsoft) cria plataformas e ferramentas robustecidas por Inteligência Artificial para entregar soluções inovadoras que correspondam às necessidades crescentes dos nossos clientes. Enquanto empresa de tecnologia está comprometida com a democratização do acesso a Inteligência Artificial de forma responsável, mantendo a missão de capacitar todas as pessoas e organizações no planeta para atingir mais.