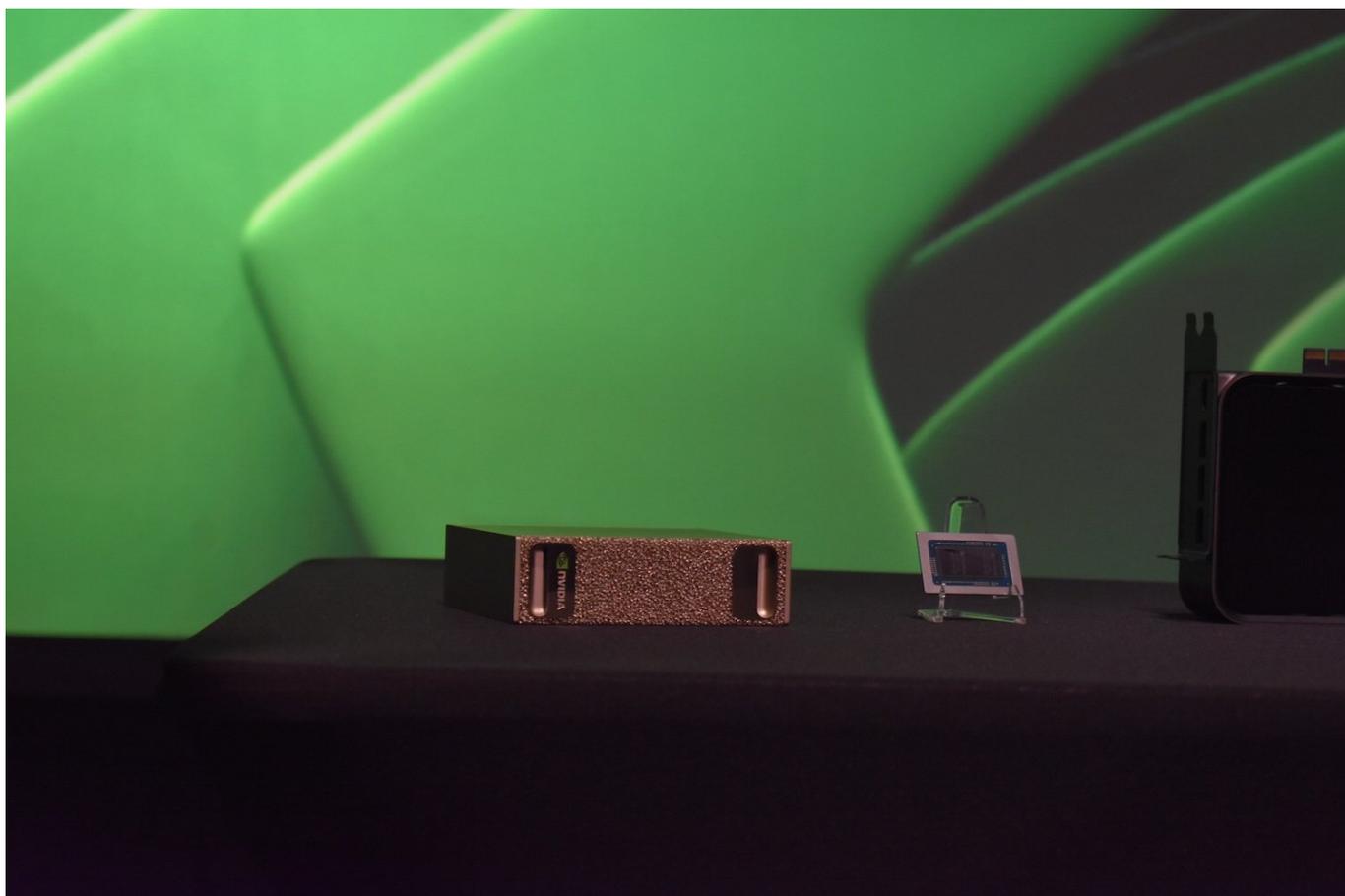


A Nvidia é uma “empresa de tecnologia”, não uma empresa de “consumidor” ou “empresarial”, como enfatizou o CEO Jensen Huang. O que ele quer dizer exatamente? A Nvidia não quer que os consumidores gastem centenas ou milhares de dólares nas novas e caras GPUs da série RTX 50? Eles não querem que mais empresas comprem seus chips de treinamento em IA? A Nvidia é o tipo de empresa que tem muitos dedos em muitas tortas. Segundo Huang, se a crosta dessas tortas são os chips da empresa, então a IA é o recheio.

“Nossa influência tecnológica terá impacto no futuro das plataformas de consumo”, disse Huang - vestido com sua típica jaqueta preta e o seio caloroso do hype da IA - em uma sessão de perguntas e respostas com repórteres um dia após sua palestra de abertura na CES. Mas como uma empresa como a Nvidia financia todos esses experimentos épicos de IA? Os chips de treinamento H100 AI fizeram da Nvidia uma potência tecnológica nos últimos dois anos, com alguns tropeços ao longo do caminho. Mas a Amazon e outras empresas estão tentando criar alternativas para acabar com o monopólio da Nvidia. O que deverá acontecer se a concorrência encurtar a onda?



O bloco Digits da Nvidia mal tem o tamanho de um mini-PC típico. © Foto: Kyle Barr/Gizmodo

“Vamos responder aos clientes onde quer que estejam”, disse Huang. Parte disso é ajudar as empresas a construir “IA agente”, também conhecida como vários modelos de IA capazes de concluir tarefas complexas. Isso inclui vários kits de ferramentas de IA feitos para ajudar as empresas. Embora o H100 tenha tornado a Nvidia grande e o RTX faça com que os



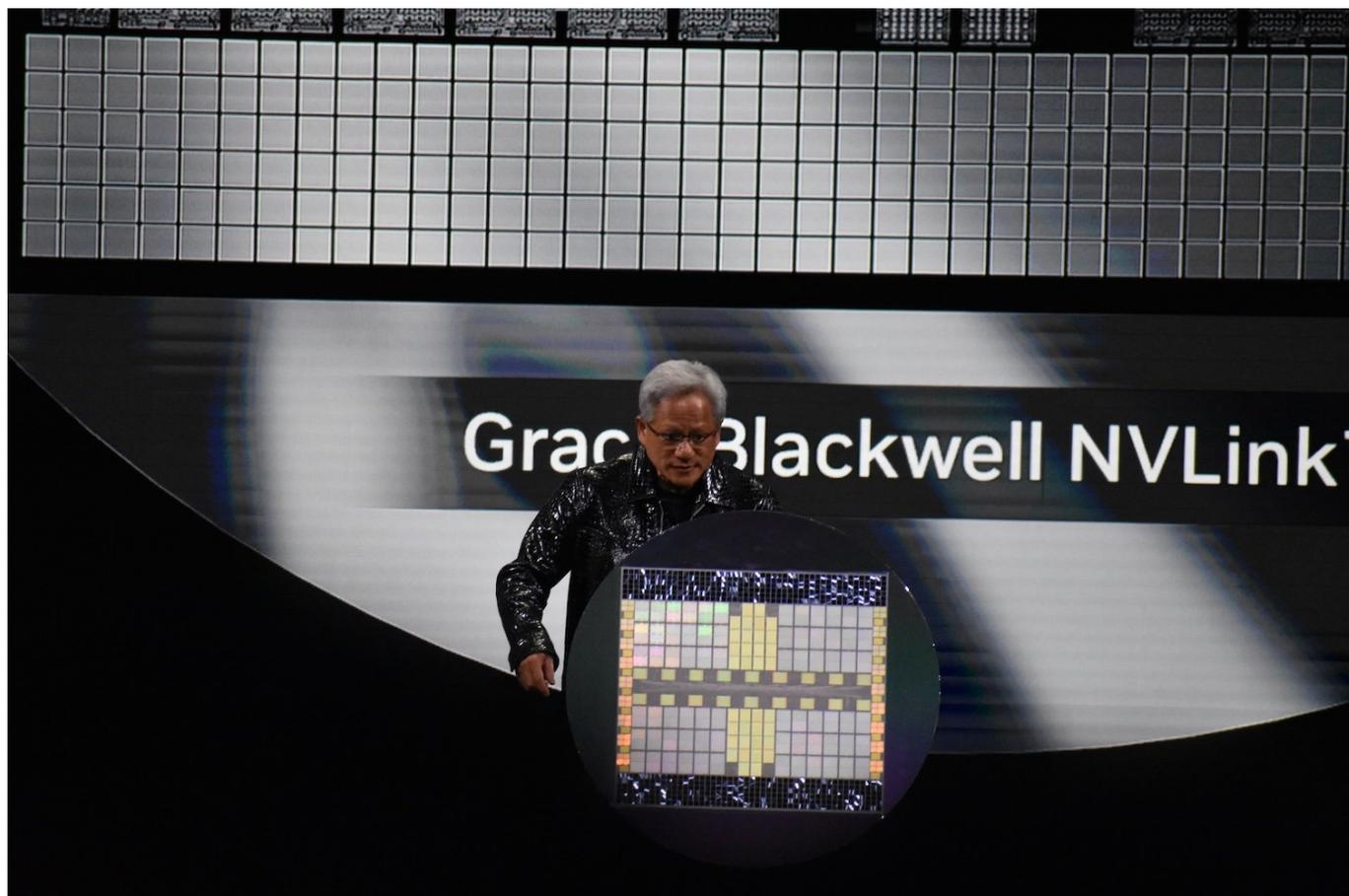
jogadores voltem, ele quer que seu novo centro de processamento de IA “Project Digits” de US\$ 3.000 abra “um universo totalmente novo” para aqueles que podem usá-lo. Quem vai usá-lo? A Nvidia disse que é uma ferramenta para pesquisadores, cientistas e talvez estudantes – ou pelo menos aqueles que tropeçam em US\$ 3.000 em sua xícara de ramen instantâneo de US\$ 1,50 que comem no jantar pela quinta noite consecutiva.

A Nvidia garantiu que você soubesse sobre os 3.352 TOPS de desempenho de IA do RTX 5090. Então, a empresa de Huang divulgou detalhes sobre várias iniciativas de software – tanto relacionadas a jogos quanto não relacionadas a jogos. Nenhuma de suas declarações foi mais confusa do que seus modelos de IA de “fundação mundial”. Esses modelos devem ser capazes de treinar em ambientes da vida real, que poderiam ser usados para ajudar veículos autônomos ou robôs a navegar em seu ambiente. É muita tecnologia do futuro, e Huang admitiu que não conseguiu articulá-la melhor para uma multidão que veio principalmente para ver novas GPUs interessantes.

“(O modelo de fundação mundial) entende coisas como fricção, inércia, agarramento, presença de objetos e elementos, compreensão geométrica e espacial”, disse ele. “Você sabe, as coisas que as crianças sabem. Eles entendem o mundo físico de uma forma que os modelos de linguagem não conheciam.”

Huang abriu a CES 2025 em 6 de janeiro com uma palestra que lotou a arena Michelob Ultra no cassino Mandalay Bay de Las Vegas. Certamente havia uma grande parte de jogadores que vieram ver as mais recentes placas da série RTX 50 em carne e osso, mas muitos estavam lá para ver como uma empresa tão lucrativa como a Nvidia avança. RTX e Project Digits atraíram gritos e gritos da multidão. Passando metade do tempo falando sobre seu modelo de fundação mundial, o público não parecia tão entusiasmado.

Isso mostra como as mensagens de IA podem ser estranhas, especialmente para uma empresa que oferece grande parte de sua popularidade à população atenta de jogadores de [PC](#). Tem-se falado tanto sobre IA que é fácil esquecer que a Nvidia estava neste jogo anos antes do [ChatGPT](#) entrar em cena. A tecnologia de upscaling de IA no jogo da Nvidia, DLSS, existe há quase seis anos, melhorando o tempo todo, e agora é um dos melhores upscalers de IA em jogos, embora limitado por sua exclusividade às placas da Nvidia. Era bom antes do advento da IA generativa. Agora, a Nvidia promete que os modelos Transformer irão melhorar ainda mais o upscaling e a reconstrução de raios.



© Foto: Kyle Barr/Gizmodo

Para completar, a alardeada geração multi-frame poderia garantir quatro vezes o desempenho para GPUs da série 50, pelo menos se o jogo suportar. Isso é uma vantagem para aqueles que podem pagar pela nova série RTX 50. O RTX 5090 custa US\$ 2.000. Os jogadores que mais se beneficiariam com a geração de quadros são aqueles que só podem pagar por uma GPU de baixo custo. Huang se recusou a oferecer qualquer dica sobre um RTX 5050 ou 5060, brincando “Anunciamos quatro placas e você quer mais?”

O modelo de fundação mundial é apenas um protótipo, assim como grande parte do novo software de IA da Nvidia em exibição ao público. As verdadeiras questões são: quando estará pronto para o horário nobre e quem acabará por usá-lo? A Nvidia exibiu NPCs de IA excêntricos, chatbots no jogo, enfermeiras de IA e um gerador de áudio no ano passado. Este ano, ele quer florescer com seu modelo de base mundial, além de uma série de “microserviços” de IA, incluindo um estranho falante animado que deveria servir como assistente sempre ativo do seu PC. Talvez alguns deles permaneçam. Nos casos em que a Nvidia espera que a IA substitua enfermeiras ou engenheiros de áudio, esperamos que isso não aconteça.

Huang considera a Nvidia “uma pequena empresa” com 32 mil funcionários em todo o mundo. Sim, isso é menos da metade da equipe que a [Meta](#) possui, mas você não pode considerá-la pequena em termos de influência no mercado de chips de treinamento de IA. Devido à sua posição no mercado, exerce uma influência descomunal na indústria de



tecnologia. Quanto mais pessoas usarem IA, mais pessoas precisarão comprar suas GPUs específicas para IA, além de qualquer outro software de IA. Se todos comprarem seu próprio chip de processamento de IA doméstico, não precisarão depender de data centers externos e chatbots externos. A Nvidia, assim como toda empresa de tecnologia, só precisa encontrar um uso para a IA além de substituir todos os nossos empregos.