Afinal, o que é memória cache?

Em resumo, a memória cache é uma parte do [processador (CPU)](https://canaltech.com.br/hardware/o-que-e-cpu/) de computadores e dispositivos móveis. Ele atua como uma memória temporária para que o chipset recupere rapidamente os dados, sem a necessidade de uma busca direta na memória principal (RAM).

Para isso, os dados de aplicativos e arquivos mais acessados pelo usuário ficam registrados nessa área de memória temporária em PCs, celulares e outros dispositivos eletrônicos. Como resultado, isso otimiza a performance dos softwares, já que o caminho até a informação é encurtado.

O termo Cache também se refere à coleta de dados temporários de softwares e de hardwares. Por isso, é sempre recomendado reiniciar os dispositivos para realizar uma limpeza de cache e permitir que os aplicativos sejam carregados mais rápido.

Por exemplo, o Google Chrome possui um cache de imagens que agilizam o carregamento das fotos ao abrir uma página da Internet. Contudo, quando essa “pasta” está lotada, o navegador fica mais lento e é necessário fazer uma limpeza dos arquivos temporários para que novos conteúdos possam ser alocados.

Memória RAM e memória cache trabalham de forma diferente com os dados temporários (Imagem: Liam Briese/Unsplash)

Cache x RAM

Alguns usuários confundem memória cache com memória RAM. Apesar de terem funções semelhantes, cada uma dessas memórias possui características diferentes.

O cache pode ser apelidado de “memória da CPU”, pois é um componente integrado ao processador ou um chip interconectado. Por estar próximo ao chipset, ele fica mais “acessível” e consegue transferir dados de forma ágil e eficiente para o componente.

Por fazer parte do processador, a memória cache é um componente menor e, por consequência, oferece menos espaço de armazenamento. Contudo, a velocidade dele é até 100 vezes maior do que a memória RAM.

Enquanto isso, a [memória RAM](https://canaltech.com.br/hardware/o-que-e-memoria-ram/) é o principal armazenamento de dados temporários de PCs e dispositivos móveis. Essa peça tem a “responsabilidade” de organizar todos os processos e aplicativos ativos, permitindo ao dispositivo processá-los rapidamente e registrá-los no HD ou SSD.

Contudo, a CPU sempre inicia a busca de dados pelo cache por conta da alta velocidade e da “proximidade” física. Então, a memória RAM segue gerenciando os demais processos de dados temporários.

É importante mencionar que o aumento de memória RAM não ampliará a memória cache. A segunda está diretamente relacionada com o processador do PC ou smartphone, é por isso não pode receber um "upgrade" para ter mais espaço para os dados temporários.

Por estar interconectada com a CPU, a memória cache tem maior velocidade de transmissão (Imagem: Roman Spiridonov/Unsplash)

Como o cache influencia no desempenho?

Como dito, o objetivo da memória cache é agilizar o acesso aos dados usados com maior frequência. Dessa maneira, o computador costuma registrar uma cópia das informações mais acessadas e guardá-la em uma área especial dessa “biblioteca de arquivos”.

Esse processo é batizado de “caching”. Quanto mais a pessoa realizar determinada ação, como abrir sempre os mesmos aplicativos, maior é a probabilidade desses dados serem copiados para o cache.

Então, o processador passa a atuar em um modo de “Acertos e Erros”. Quando a pessoa executa uma tarefa, o chipset vai procurar os dados no cache e, caso encontre o que busca, isso será sinalizado como um “acerto” e deixa o desempenho do dispositivo mais rápido.

Se os dados não estiverem na memória cache, isso é sinalizado como um “erro”. Então, o processador terá que fazer um caminho mais longo – e microsegundos mais “demorado” – até encontrar as informações na memória RAM ou no armazenamento interno.

Por ser uma memória mais limitada, deixar o cache cheio de cópias causa lentidão. Limpar os arquivos temporários do dispositivo e dos apps instalados é o segredo para liberar mais espaço e manter o alto desempenho.

"Hierarquia" da memória cache e a relação com a memória RAM e armazenamento interno (Imagem: Reprodução/Modern Hardware)

Categorias de cache

É importante destacar que a memória cache é uma “grande biblioteca” de dados temporários. Contudo, essas informações de rápido acesso costumam ser organizadas em níveis de prioridade que otimizam o trabalho do processador.

Cache 1 (L1)

Mais próximo ao chipset, o Cache 1 – ou apenas, L1 – avalia os dados acessados recentemente pelo usuário. Assim, a função dele é determinar quais informações têm maior probabilidade de serem acessadas novamente.

Esse é o primeiro espaço de armazenamento em que o processador busca os dados quando precisa encontrar certas informações. Embora seja o nível de cache com maior velocidade, ele é relativamente menor que os demais.

Cache 2 (L2)

O Cache 2 (L2), ou “cache secundário”, é o espaço de armazenamento em que o processador realiza a segunda busca por dados após não encontrá-los no L1. Aqui são armazenadas as informações acessadas em momentos mais recentes.

Essa área costuma ter velocidade e tamanho médio devido à “distância” em que ela está do processador. Contudo, esse espaço é interligado com o L1 e os dados requisitados podem ser transmitidos para o primeiro nível da memória cache.

Cache 3 (L3)

A função do Cache 3 (L3) é auxiliar no desempenho do L1 e do L2. Embora seja o maior espaço e mais lento dos três níveis, ele ainda atua em uma velocidade superior à memória RAM do dispositivo.

Em processadores de múltiplos núcleos, cada núcleo pode ter os caches L1 e L2 dedicados e compartilharem o L3. Se o terceiro cache tiver os dados buscados pelo computador, essas informações são transmitidas para o nível mais alto.

Cache de Disco

Como o nome adianta, o Cache de Disco é a memória de dados temporários relacionado ao armazenamento interno do computador ou dispositivo móvel. Esse costuma ser o nível de cache mais lento de todos, pois busca todos os dados registrados na memória RAM.

Vale mencionar que a memória RAM também organiza informações temporárias de periféricos conectados ao PC. Por exemplo, os dados de um HD externo ficarão registrados em um cache especial para esse acessório.

Placas de vídeo dedicadas têm memoria cache independente da CPU do computador (Imagem: Reprodução/Tech Blogs)

Cache da GPU

Para que a renderização de imagens aconteça com maior velocidade, os computadores possuem um cache dedicado para os dados temporários da placa gráfica. Esse processo pode acontecer de forma diferente com GPUs integradas ou placas dedicadas.

No caso das placas de vídeos integradas ao processador, as unidades compartilham a memória cache. Por consequência, o armazenamento de dados temporários da GPU é mais limitado por conta da disputa de “espaço” com outras informações.

Já as placas gráficas dedicadas, separadas da CPU, têm o próprio armazenamento de dados temporários. Em PCs gamer, a vantagem do cache independente é redução de atrasos ou travamentos em títulos que exigem um maior ritmo de renderização.

Conseguiu entender os segredos da memória cache? Não esqueça de compartilhar esse conteúdo com outras pessoas interessadas no tema!