

How can I get the memory manager to prefetch bigger chunks of data from my memory-mapped file?

 devblogs.microsoft.com/oldnewthing/20160909-00

September 9, 2016



Raymond Chen

A customer had a memory-mapped file and they found that when they accessed a page in the mapping, Windows paged in 32KB of memory, even though the documentation says that only 4KB gets paged in. The customer's application reads 100 small records from a memory-mapped file on an SSD, so latency is the most important factor. They were hoping for a way to get the prefetcher to prefetch bigger chunks of the memory-mapped file.

Okay, let's take things one issue at a time.

Why are they observing 32KB reads when the documentation says 4KB? The operating system's contractual obligation is to bring in the entire page, which is 4KB on x86-class machines. However, the operating system is allowed to perform speculative reads, and Windows will read up to 32KB of memory around the faulting page. The precise amount depends on a variety of factors, including how the memory was mapped, which pages are already present, and (for pagefile-backed memory), whether the pages are contiguous in the pagefile.¹

What the customer can do is call `PrefetchVirtualMemory` to initiate explicit prefetching.

The customer wrote back that with the explicit call to `PrefetchVirtualMemory`, the I/O system sends all the requests to the device at once, "which seems to be exactly what we need."

¹The maximum automatic prefetch for pagefile-backed memory is 64KB, but this increase is not as big a deal as it sounds, because in practice, consecutive addresses in memory tend not to be assigned to consecutive pages in the pagefile, so the speculative read from the pagefile tends not to read very much.

[Raymond Chen](#)

Follow

