What order are pages in a memory-mapped file flushed out to disk?

devblogs.microsoft.com/oldnewthing/20190307-00

March 7, 2019

Raymond Chen

Say you create a memory-mapped file mapping and write data to it. In which order will the pages be flushed out to disk? Is it in the order the pages were written? Or are pages closer to the beginning of the file flushed out before pages closer to the end of the file?

There is no defined order in which the pages are written to disk.

In practice, it's a combination of multiple factors, many of which you don't really have control over. Tim Sneath goes into the details from the memory manager's point of view, along with a nice picture. To answer the question of when the pages get written to disk, we need to follow the flow up to the point where the memory manager decides to write the dirty page to disk.

When the page falls out of the working set of every process that has it mapped, then it goes to the modified page list. And it is when the page reaches the modified page list that it gets written to disk (if dirty) before being moved to the standby list. So the order in which the memory manager writes the pages to disk is the order in which they fall out of the working set of every process that has it mapped.

But wait, we're not finished yet. That's the order that the memory manager writes the pages, but that's not necessarily the order in which the pages make it to disk. That's merely the order in which the memory manager submits the pages to the I/O subsystem.

The I/O subsystem will reorder the requests in order to complete them efficiently. For example, if there are requests for hard drive sectors 12, 80, and 13, the driver will probably write sector 12, then 13, then 80 in order to minimize drive head travel time.

And that's assuming that you have rotational media.

If you have an SSD, then drive head travel time is not a factor, but the block size enters the picture. Sectors within the same block are likely to be grouped together, so you are likely to see sectors 12 and 13 written simultaneously, and sector 80 written as a separate command.

The order in which pages in a memory-mapped file are flushed out to disk is therefore not only not contractual, but also unpredictable.

Raymond Chen

Follow