

Direct Memory Access (DMA)

Direct Memory Access is a method of allowing data to be moved from one location to another in a computer without intervention from the central processor (CPU). The way that the DMA function is implemented varies between computer architectures, so this discussion will limit itself to the implementation and workings of the DMA subsystem on the IBM Personal Computer (PC), the IBM PC/AT and all of its successors and clones. The PC DMA subsystem is based on the Intel® 8237 DMA controller. The 8237 contains four DMA channels that can be programmed independently and any one of the channels may be active at any moment. These channels are numbered 0, 1, 2 and 3. Starting with the PC/AT, IBM added a second 8237 chip, and numbered those channels 4, 5, 6 and 7. The original DMA controller (0, 1, 2 and 3) moves one byte in each transfer. The second DMA controller (4, 5, 6, and 7) moves 16-bits from two adjacent memory locations in each transfer, with the first byte always coming from an even-numbered address. The two controllers are identical components and the difference in transfer size is caused by the way the second controller is wired into the system. The 8237 has two electrical signals for each channel, named DRQ and -DACK. There are additional signals with the names HRQ (Hold Request), HLDA (Hold Acknowledge), -EOP (End of Process), and the bus control signals -MEMR (Memory Read), -MEMW (Memory Write), -IOR (I/O Read), and -IOW (I/O Write). The 8237 DMA is known as a “fly-by” DMA controller. This means that the data being moved from one location to another does not pass through the DMA chip and is not stored in the DMA chip. Subsequently, the DMA can only transfer data between an I/O port and a memory address, but not between two I/O ports or two memory locations.

