# 2. API同步行为（synchronization behavior）

**Memcpy**

API以同步和异步的形式提供memcpy/memset函数，后者具有“Async”后缀。这是一个用词不当的词，因为每个函数都可能表现出同步或异步行为，这取决于传递给函数的参数。

**同步（Synchronous）**

1. 涉及统一内存区域（Unifined Memory regions）的所有传输相对于主机完全同步。
2. 从可分页主机内存（pageable host memory）向设备内存传输，在复制完成之前会执行流同步。对于DMA传输到设备内存，一旦将可分页缓冲区（pageable buffer）复制到暂存区（staging memory），该函数将返回（此时DMA传输到目的地址可能未完成）。
3. 从固定主机内存（pinned host memory）向设备传输，该函数相对于主机同步。
4. 从设备向主机的可分页内存或固定内存（pageable or pinned host memory）传输，改函数直到复制完成之后才返回。
5. **设备内的内存复制，主机端不参与同步。**
6. 主机内的内存复制，相对主机执行完全同步。

**异步（Asynchronous，即cudaMemcpyAsync()函数）**

1. 从设备向主机可分页内存传输，函数直到复制完成之后才返回。
2. 主机间的任何内存复制，相对主机完全同步。
3. 除此之外，该函数是异步的。如果可分页内存必须首先被暂存到固定内存，将使用辅助线程（a worker thread）异步处理。

**Memset**

同步memset函数相对于主机而言是异步的，**除非目标是固定主机内存或统一内存区域中**，在这种情况下，它们是完全同步的。

异步版本对于主机始终是异步的。

**内核启动（Kernel Launches）**

相对于主机，内核启动是异步的。