

**为什么我们用netperf测不出来预期的结果？文中提到**

Under regular circumstances, the IOTLB has not been observed to be a bottleneck so far. For several devices, the virtual I/O memory **map** and **unmap** operations consume CPU-time, which is greater than the time of the corresponding DMA transaction.

As demonstrated in Section 2, once the computational cost of frequent mapping and unmapping of IOMMU buffers is sufficiently reduced [19, 20], the address resolution mechanism becomes the main bottleneck.

由于我们用netperf测throughput是DMA的过程，**map**和**unmap**操作所花的时间比地址解析时间可能大了不止一个数量级，因而当IOTLB被刷新时，即使地址解析的latency增大，我们仍然观察不出throughput受影响的结果。

**文中的解决方案：**如果要观察到IOTLB的overhead，按文中所说：

First, to eliminate the time required by the CPU for DMA map and unmap operations, we used the pseudo pass-through mode of the IOMMU. This mode works by using a fixed identity mapping in the IOMMU page tables for all of the devices, thus eliminating most of the mapping and unmapping operations and their associated overhead. In addition, this mode uses static mappings— it does not change mappings and does not flush IOTLB entries as a use-once mapping strategy does.

Second, we constructed a stress-test micro-benchmark using a high-speed I/O device. For this purpose our experiments utilized Intel’s I/O Acceleration Technology (I/OAT) which enables asynchronous DMA memory copy in bandwidths of over 16Gbps per channel [18].

另： 网卡在做DMA操作时，有两种DMA映射：

A. Consistent DMA mappings (sometimes referred to as Coherent DMA Mappings) are persistent mappings that are usually mapped once at driver initialization and unmapped when the driver is removed. These mappings are used by network cards for DMA ring descriptors.

Note: 这种映射只针对与网卡相关的必要数据结构描述符，而不是DMA操作的数据包部分。

B. Streaming DMA mappings are ephemeral mappings that are usually mapped for one DMA transaction and unmapped as soon as it completes. Such mapping are used for networking buffers transmitted or received by NICs and for file-system buffers written or read by a SCSI device.

Note: 这种映射就是所熟悉的外设做DMA操作时所涉及到I/O数据包部分。

**因此，**我们需要修改xen中IOMMU对外设的映射方式，使得每次DMA操作时的I/O data buffer的映射关系均没有发生变化，变成静态映射，比如说：

Permanently map the entire physical address space of the guest operating system. Such a strategy is sometimes referred to as a direct map, because this arrangement creates a one-to-one mapping between IOMMU entries and physical pages for each physical page owned by the guest operating system.

该段出自：**Protection Strategies for Direct Access to Virtualized I/O Devices**

减少map和unmap的时间开销；通过增加buffer大小，促使IOTLB cache达到阈值，由此观察时间花销。