在各类电子设备和元器件中，我们都可以接触到带宽的概念，例如我们熟知的显示器的带宽、内存的带宽、总线的带宽和网络的带宽等等；对这些设备而言，带宽是一个非常重要的指标。不过容易让人迷惑的是，在显示器中它的单位是MHz，这是一个频率的概念；而在总线和内存中的单位则是GB/s，相当于数据传输率的概念；而在通讯领域，带宽的描述单位又变成了MHz、GHz……这两种不同单位的带宽表达的是同一个内涵么？二者存在哪些方面的联系呢？本文就带你走入精彩的带宽世界。  
一、 带宽的两种概念  
如果从电子电路角度出发，带宽（Bandwidth）本意指的是电子电路中存在一个固有通频带，这个概念或许比较抽象，我们有必要作进一步解释。大家都知道，各类复杂的电子电路无一例外都存在电感、电容或相当功能的储能元件，即使没有采用现成的[电感线圈](https://www.baidu.com/s?wd=%E7%94%B5%E6%84%9F%E7%BA%BF%E5%9C%88&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd)或电容，导线自身就是一个电感，而导线与导线之间、导线与地之间便可以组成电容——这就是通常所说的杂散电容或分布电容；不管是哪种类型的电容、电感，都会对信号起着阻滞作用从而消耗信号能量，严重的话会影响信号品质。这种效应与交流电信号的频率成正比关系，当频率高到一定程度、令信号难以保持稳定时，整个电子电路自然就无法正常工作。为此，电子学上就提出了“带宽”的概念，它指的是电路可以保持稳定工作的频率范围。而属于该体系的有显示器带宽、通讯/网络中的带宽等等。  
而第二种带宽的概念大家也许会更熟悉，它所指的其实是数据传输率，譬如内存带宽、[总线带宽](https://www.baidu.com/s?wd=%E6%80%BB%E7%BA%BF%E5%B8%A6%E5%AE%BD&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd)、网络带宽等等，都是以“字节/秒”为单位。我们不清楚从什么时候起这些数据传输率的概念被称为“带宽”，但因业界与公众都接受了这种说法，代表数据传输率的带宽概念非常流行，尽管它与电子电路中“带宽”的本意相差很远。  
对于电子电路中的带宽，决定因素在于[电路设计](https://www.baidu.com/s?wd=%E7%94%B5%E8%B7%AF%E8%AE%BE%E8%AE%A1&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd)。它主要是由高频放大部分元件的特性决定，而高频电路的设计是比较困难的部分，成本也比普通电路要高很多。这部分内容涉及到[电路设计](https://www.baidu.com/s?wd=%E7%94%B5%E8%B7%AF%E8%AE%BE%E8%AE%A1&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd)的知识，对此我们就不做深入的分析。而对于总线、内存中的带宽，决定其数值的主要因素在于工作频率和位宽，在这两个领域，带宽等于工作频率与位宽的乘积，因此带宽和工作频率、位宽两个指标成正比。不过工作频率或位宽并不能无限制提高，它们受到很多因素的制约，我们会在接下来的总线、内存部分对其作专门论述。  
二、 总线中的带宽  
在计算机系统中，总线的作用就好比是人体中的神经系统，它承担的是所有数据传输的职责，而各个子系统间都必须籍由总线才能通讯，例如，CPU和北桥间有前端总线、北桥与显卡间为AGP总线、芯片组间有南北桥总线，各类扩展设备通过PCI、[PCI-X](https://www.baidu.com/s?wd=PCI-X&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd" \t "_blank)总线与系统连接；主机与外部设备的连接也是通过总线进行，如目前流行的[USB](https://www.baidu.com/s?wd=USB&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd) [2.0](https://www.baidu.com/s?wd=2.0&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd)、IEEE1394总线等等，一句话，在一部计算机系统内，所有数据交换的需求都必须通过总线来实现！  
按照工作模式不同，总线可分为两种类型，一种是并行总线，它在同一时刻可以传输多位数据，好比是一条允许多辆车并排开的宽敞道路，而且它还有双向单向之分；另一种为串行总线，它在同一时刻只能传输一个数据，好比只容许一辆车行走的狭窄道路，数据必须一个接一个传输、看起来仿佛一个长长的数据串，故称为“串行”。  
并行总线和串行总线的描述参数存在一定差别。对并行总线来说，描述的性能参数有以下三个：总线宽度、[时钟频率](https://www.baidu.com/s?wd=%E6%97%B6%E9%92%9F%E9%A2%91%E7%8E%87&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd)、数据传输频率。其中，总线宽度就是该总线可同时传输数据的位数，好比是车道容许并排行走的车辆的数量；例如，16位总线在同一时刻传输的数据为16位，也就是2个字节；而32位总线可同时传输4个字节，64位总线可以同时传输8个字节......显然，总线的宽度越大，它在同一时刻就能够传输更多的数据。不过总线的位宽无法无限制增加。[时钟频率](https://www.baidu.com/s?wd=%E6%97%B6%E9%92%9F%E9%A2%91%E7%8E%87&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd)和数据传输频率的概念在上一期的文章中有过详细介绍，我们就不作赘述。  
总线的带宽指的是这条总线在单位时间内可以传输的数据总量，它等于总线位宽与工作频率的乘积。例如，对于64位、800MHz的前端总线，它的数据传输率就等于64bit×800MHz÷8(Byte)=6.4GB/s；32位、33MHz PCI总线的数据传输率就是32bit×33MHz÷8=133MB/s，等等，这项法则可以用于所有并行总线上面——看到这里，读者应该明白我们所说的[总线带宽](https://www.baidu.com/s?wd=%E6%80%BB%E7%BA%BF%E5%B8%A6%E5%AE%BD&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd)指的就是它的数据传输率，其实“总线带宽”的概念同“电路带宽”的原始概念已经风马牛不相及。  
对串行总线来说，带宽和工作频率的概念与并行总线完全相同，只是它改变了传统意义上的总线位宽的概念。在频率相同的情况下，并行总线比串行总线快得多，那么，为什么现在各类并行总线反而要被串行总线接替呢？原因在于并行总线虽然一次可以传输多位数据，但它存在并行传输信号间的干扰现象，频率越高、位宽越大，干扰就越严重，因此要大幅提高现有并行总线的带宽是非常困难的；而串行总线不存在这个问题，总线频率可以大幅向上提升，这样串行总线就可以凭借高频率的优势获得高带宽。而为了弥补一次只能传送一位数据的不足，串行总线常常采用多条管线（或通道）的做法实现更高的速度——管线之间各自独立，多条管线组成一条总线系统，从表面看来它和并行总线很类似，但在内部它是以串行原理运作的。对这类总线，带宽的计算公式就等于“总线频率×管线数”，这方面的例子有PCI Express和HyperTransport，前者有×1、×2、×4、×8、×16和×32多个版本，在第一代PCI Express技术当中，单通道的单向信号频率可达2.5GHz，我们以×16举例，这里的16就代表16对双向总线，一共64条线路，每4条线路组成一个通道，二条接收，二条发送。这样我们可以换算出其总线的带宽为2.5GHz×16/10=4GB/s（单向）。除10是因为每字节采用10位编码。  
三、 内存中的带宽  
除总线之外，内存也存在类似的带宽概念。其实所谓的内存带宽，指的也就是内存总线所能提供的数据传输能力，但它决定于[内存芯片](https://www.baidu.com/s?wd=%E5%86%85%E5%AD%98%E8%8A%AF%E7%89%87&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd)和内存模组而非纯粹的总线设计，加上地位重要，往往作为单独的对象讨论。  
SDRAM、DDR和DDRⅡ的总线位宽为64位，RDRAM的位宽为16位。而这两者在结构上有很大区别：SDRAM、DDR和DDRⅡ的64位总线必须由多枚芯片共同实现，计算方法如下：内存模组位宽=[内存芯片](https://www.baidu.com/s?wd=%E5%86%85%E5%AD%98%E8%8A%AF%E7%89%87&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd)位宽×单面芯片数量（假定为单面单物理BANK）；如果[内存芯片](https://www.baidu.com/s?wd=%E5%86%85%E5%AD%98%E8%8A%AF%E7%89%87&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd)的位宽为8位，那么模组中必须、也只能有8颗芯片，多一枚、少一枚都是不允许的；如果芯片的位宽为4位，模组就必须有16颗芯片才行，显然，为实现更高的模组容量，采用高位宽的芯片是一个好办法。而对RDRAM来说就不是如此，它的内存总线为串联架构，总线位宽就等于内存芯片的位宽。  
和并行总线一样，内存的带宽等于位宽与数据传输频率的乘积，例如，[DDR400](https://www.baidu.com/s?wd=DDR400&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd)内存的数据传输频率为400MHz，那么单条模组就拥有64bit×400MHz÷8(Byte)=3.2GB/s的带宽；PC 800标准RDRAM的频率达到800MHz，单条模组带宽为16bit×800MHz÷ 8=1.6GB/s。为了实现更高的带宽，在[内存控制器](https://www.baidu.com/s?wd=%E5%86%85%E5%AD%98%E6%8E%A7%E5%88%B6%E5%99%A8&tn=44039180_cpr&fenlei=mv6quAkxTZn0IZRqIHckPjm4nH00T1Ykm1uWmhmzuWf1nWc1PjT40ZwV5Hcvrjm3rH6sPfKWUMw85HfYnjn4nH6sgvPsT6KdThsqpZwYTjCEQLGCpyw9Uz4Bmy-bIi4WUvYETgN-TLwGUv3EPHD3PjmzPj6YPjnvnHb4rHcd)中使用双通道技术是一个理想的办法，所谓双通道就是让两组内存并行运作，内存的总位宽提高一倍，带宽也随之提高了一倍！  
带宽可以说是内存性能最主要的标志，业界也以内存带宽作为主要的分类标准，但它并非决定性能的唯一要素，在实际应用中，内存延迟的影响并不亚于带宽。如果延迟时间太长的话相当不利，此时即便带宽再高也无济于事。  
四、 带宽匹配的问题  
计算机系统中存在形形色色的总线，这不可避免带来总线速度匹配问题，其中最常出问题的地方在于前端总线和内存、南北桥总线和PCI总线。  
前端总线与内存匹配与否对整套系统影响最大，最理想的情况是前端总线带宽与内存带宽相等，而且内存延迟要尽可能低。在Pentium4刚推出的时候，Intel采用RDRAM内存以达到同前端总线匹配，但RDRAM成本昂贵，严重影响推广工作，Intel曾推出搭配PC133 SDRAM的845芯片组，但SDRAM仅能提供1.06GB/s的带宽，仅相当于400MHz前端总线带宽的1/3，严重不匹配导致系统性能大幅度下降；后来，Intel推出支持DDR266的845D才勉强好转，但仍未实现与前端总线匹配；接着，Intel将P4前端总线提升到533MHz、带宽增长至5.4GB/s，虽然配套芯片组可支持DDR333内存，可也仅能满足1/2而已；现在，P4的前端总线提升到800MHz，而配套的865/875P芯片组可支持双通道DDR400——这个时候才实现匹配的理想状态，当然，这个时候继续提高内存带宽意义就不是特别大，因为它超出了前端总线的接收能力。  
南北桥总线带宽曾是一个尖锐的问题，早期的芯片组都是通过PCI总线来连接南北桥，而它所能提供的带宽仅仅只有133MB/s，若南桥连接两个ATA-100硬盘、100M网络、IEEE1394接口......区区133MB/s带宽势必形成严重的瓶颈，为此，各芯片组厂商都发展出不同的南北桥总线方案，如Intel的Hub-Link、VIA的V-Link、SiS 的MuTIOL，还有AMD的 HyperTransport等等，目前它们的带宽都大大超过了133MB/s，最高纪录已超过1GB/s，瓶颈效应已不复存在。  
PCI总线带宽不足还是比较大的矛盾，目前PC上使用的PCI总线均为32位、33MHz类型，带宽133MB/s，而这区区133MB/s必须满足网络、硬盘控制卡（如果有的话）之类的扩展需要，一旦使用千兆网络，瓶颈马上出现，业界打算自2004年开始以PCI Express总线来全面取代PCI总线，届时PCI带宽不足的问题将成为历史。  
五、 显示器中的带宽  
以上我们所说的“带宽”指的都是速度概念，但对CRT显示器来说，它所指的带宽则是频率概念、属于电路范畴，更符合“带宽”本来的含义。  
要了解显示器带宽的真正含义，必须简单介绍一下CRT显示器的工作原理——由灯丝、阴极、控制栅组成的电子枪，向外发射电子流，这些电子流被拥有高电压的加速器加速后获得很高的速度，接着这些高速电子流经过透镜聚焦成极细的电子束打在屏幕的荧光粉层上，而被电子束击中的地方就会产生一个光点；光点的位置由偏转线圈产生的磁场控制，而通过控制电子束的强弱和通断状态就可以在屏幕上形成不同颜色、不同灰度的光点——在某一个特定的时刻，整个屏幕上其实只有一个点可以被电子束击中并发光。为了实现满屏幕显示，这些电子束必须从左到右、从上到下一个一个象素点进行扫描，若要完成800×600分辨率的画面显示，电子枪必须完成800×600=480000个点的顺序扫描。由于荧光粉受到电子束击打后发光的时间很短，电子束在扫描完一个屏幕后必须立刻再从头开始——这个过程其实十分短暂，在一秒钟时间电子束往往都能完成超过85个完整画面的扫描、屏幕画面更新85次，人眼无法感知到如此小的时间差异会“误以为”屏幕处于始终发亮的状态。而每秒钟屏幕画面刷新的次数就叫场频，或称为屏幕的垂直扫描频率、以Hz（赫兹）为单位，也就是我们俗称的“刷新率”。以800×600分辨率、85Hz刷新率计算，电子枪在一秒钟至少要扫描800×600×85=40800000个点的显示；如果将分辨率提高到1024×768，将刷新率提高到100Hz，电子枪要扫描的点数将大幅提高。  
按照业界公认的计算方法，显示器带宽指的就是显示器的电子枪在一秒钟内可扫描的最高点数总和，它等于“水平分辨率×垂直分辨率×场频（画面刷新次数）”，单位为MHz(兆赫)；由于显像管电子束的扫描过程是非线性的，为避免信号在扫描边缘出现衰减影响效果、保证图像的清晰度，总是将边缘扫描部分忽略掉，但在电路中它们依然是存在的。因此，我们在计算显示器带宽的时候还应该除一个取值为0.6~0.8 的“有效扫描系数”，故得出带宽计算公式如下：“带宽＝水平像素（行数）×垂直像素（列数）×场频（刷新频率）÷扫描系数”。扫描系数一般取为0.744。例如，要获得分辨率1024×768、刷新率85Hz的画面，所需要的带宽应该等于：1024×768×85÷0.744，结果大约是90MHz。  
不过，这个定义并不符合带宽的原意，称之为“像素扫描频率”似乎更为贴切。带宽的 最初概念确实也是电路中的问题——简单点说就是：在“带宽”这个频率宽度之内，放大器可以处于良好的工作状态，如果超出带宽范围，信号会很快出现衰减失真现象。从本质上说，显示器的带宽描述的也是控制电路的频率范围，带宽高低直接决定显示器所能达到的性能等级。由于前文描述的“像素扫描频率”与控制电路的“带宽”基本是成正比关系，显示器厂商就干脆把它当作显示器的“带宽”——这种做法当然没有什么错，只是容易让人产生认识上的误区。当然，从用户的角度考虑没必要追究这么多，毕竟以“像素扫描频率”作为“带宽”是很合乎人们习惯的，大家可方便使用公式计算出达到某种显示状态需要的最低带宽数值。  
但是反过来说，“带宽数值完全决定着屏幕的显示状态”是否也成立呢？答案是不完全成立，因为屏幕的显示状态除了与带宽有关系之外，还与一个重要的概念相关——它就是“行频”。行频又称为“水平扫描频率”，它指的是电子枪每秒在荧光屏上扫描过的水平线数量，计算公式为：“行频＝垂直分辨率×场频（画面刷新率）×1.07”，其中1.07为校正参数，因为显示屏上下方都存在我们看不到的区域。可见，行频是一个综合分辨率和刷新率的参数，行频越大，显示器就可以提供越高的分辨率或者刷新率。例如，1台17寸显示器要在1600×1200分辨率下达到75Hz的刷新率，那么带宽值至少需要221MHz，行频则需要96KHz，两项条件缺一不可；要达到这么高的带宽相对容易，而要达到如此高的行频就相当困难，后者成为主要的制约因素，而出于商业因素考虑，显示器厂商会突出带宽而忽略行频，这种宣传其实是一种误导。  
六、 通讯中的带宽  
在通讯和网络领域，带宽的含义又与上述定义存在差异，它指的是网络信号可使用的最高频率与最低频率之差、或者说是“频带的宽度”，也就是所谓的“Bandwidth”、“信道带宽”——这也是最严谨的技术定义。  
在100M以太网之类的铜介质布线系统中，双绞线的信道带宽通常用MHz为单位，它指的是信噪比恒定的情况下允许的信道频率范围，不过，网络的信道带宽与它的数据传输能力（单位Byte/s）存在一个稳定的基本关系。我们也可以用高速公路来作比喻：在高速路上，它所能承受的最大交通流量就相当于网络的数据运输能力，而这条高速路允许形成的宽度就相当于网络的带宽。显然，带宽越高、数据传输可利用的资源就越多，因而能达到越高的速度；除此之外，我们还可以通过改善信号质量和消除瓶颈效应实现更高的传输速度。  
网络带宽与数据传输能力的正比关系最早是由贝尔实验室的工程师Claude Shannon所发现，因此这一规律也被称为Shannon定律。而通俗起见普遍也将网络的数据传输能力与“网络带宽”完全等同起来，这样“网络带宽”表面上看与“总线带宽”形成概念上的统一，但这两者本质上就不是一个意思、相差甚远。  
七、 总结：带宽与性能  
对总线和内存来说，带宽高低对系统性能有着举足轻重的影响——倘若总线、内存的带宽不够高的话，处理器的工作频率再高也无济于事，因此带宽可谓是与频率并立的两大性能决定要素。而对CRT显示器而言，带宽越高，往往可以获得更高的分辨率、显示精度越高，不过现在CRT显示器的带宽都能够满足标准分辨率下85Hz刷新率或以上的显示需要（相信没有太多的朋友喜欢用非常高的分辨率去运行程序或者游戏），这样带宽高低就不是一个太敏感的参数了，当然，如果你追求高显示品质那是另一回事了