

学习总结（一）：深度学习实验环境和硬件配置知识总结

转载

今天又是充满希望的一天  于 2020-05-11 22:20:55 发布  5038  收藏 23

分类专栏: [debug](#) [计算机基础](#) 文章标签: [深度学习](#) [人工智能](#) [深度学习主机](#) [python](#)

原文链接: <https://www.zhihu.com/question/33996159/answer/698980270>

版权



[debug](#) 同时被 2 个专栏收录

47 篇文章 0 订阅

订阅专栏



[计算机基础](#)

13 篇文章 10 订阅

订阅专栏

实验平台

需求

现在的疫情很让人捉急：耽误了很多的时间、在家效率慢、实验室搞的云服务器每次要重新装环境等等问题，我都快吐了。自己想乘着这个机会配置一个2080ti主机，但是这几天了解到这个配置得1.7w左右，这个价格真的让人心痛，要是有1w2左右我就买了。这里附带一个咸鱼卖家给我的2070s配置清单。

```
cpu:r7 3700
主板:微星b450迫击炮 max
散热:玄冰400
显卡:rtx2070s七彩虹三风扇
电源:长城650w额定 金牌直出
内存:2x8 16gb ddr4 3000 威刚游戏威龙
硬盘:m.2 西数sn550 512gb
机箱:爱国者m2白色
```

你以后想升到或者直接用到2080s

cpu至少r7 3700

<https://blog.csdn.net/liupeng19970119>

上图这个配置清单预算是8000+，心动的感觉有没有，买了它生产力直接翻倍。每次到晚上就点开看看，胡思乱想，但是白天就被理性给压住了。买2080s型号的显卡后自己大多数实验都可以跑，而且不用每次搭建环境，这大大提高了我的生产力。我现在买了google pro，尽管显卡的问题解决了，但文件的速度却慢了下来，空空如也的显卡却让人觉得力不从心。自己有一个比长远的目标也吃硬件的配置。等一下，我怎么感觉自己为自己花钱找了好多理由！先打住。

基本知识

以前自己对计算机硬件了解可以说是小白，这些天查了许多的资料，了解了一些大概的知识。这里我来记一下基本常识，主要参照知乎大佬的文章。

<https://www.zhihu.com/question/33996159/answer/698980270>

硬件设备，是任何一名深度学习er不可或缺的核心装备。各位初级调参魔法师们，你们有没有感到缺少那一根命中注定的魔杖呢？



显卡

(GPU) 是深度学习的重要部件，甚至比CPU更重要。做深度学习不用GPU只用CPU显然是不明智的。



选购GPU有三大注意事项：性价比、显存、散热。使用16bit的RTX 2070或者RTX 2080 Ti性价比更高。除了GPU核心，显存（GPU RAM）也是不可忽视的部分。RTX比GTX系列显卡在显存方面更具优势，在显存相同的情况下，RTX能够训练两倍大的模型。通常对显存的要求如下：

如果想在研究中追求最高的成绩：显存 \geq 11 GB；

在研究中搜寻有趣新架构：显存 \geq 8 GB；

其他研究：8GB；

Kaggle竞赛：4~8GB；

创业公司：8GB（取决于具体应用的模型大小）打造网络原型8GB，训练不小于11GB

需要注意的是，如果你购买了多个RTX显卡，一定不要忽视散热。两个显卡堆叠在相邻PCIe插槽，很容易令GPU过热降频，可能导致性能下降30%。

内存

选择内存时候有两个参数：频率和容量。其中频率作用基本没有。而内存大小不会影响深度学习性能，但是它可能会影响你执行GPU代码的效率。内存容量大一点，CPU就可以不通过磁盘，直接和GPU交换数据。所以用户应该配备与GPU显存匹配的内存容量。如果有一个24GB显存的Titan RTX，应该至少有24GB的内存。但是，如果有更多的GPU，则不一定需要更多内存。Tim认为：内存关系到你能不能集中资源，解决更困难的编程问题。如果有更多的内存，你就可以将注意力集中在更紧迫的问题上，而不用花大量时间解决内存瓶颈。他还在参加Kaggle比赛的过程中发现，额外的内存对特征工程非常有用。这一点怎么理解呢？GPU可以直接和内存交互数据，如果内存比显存还大的话，还能做一些额外的工作，如特征工程。

CPU

过分关注CPU的性能和PCIe通道数量，是常见的认知误区。用户更需要关注的是CPU和主板组合支持同时运行的GPU数量。

CPU和PCIe人们对PCIe通道的执念近乎疯狂！而实际上，它对深度学习性能几乎没有影响。如果只有一个GPU，PCIe通道的作用只是快速地将数据从内存传输到显存。ImageNet里的32张图像（32x225x225x3）在16通道上传输需要1.1毫秒，在8通道上需要2.3毫秒，在4通道上需要4.5毫秒。这些只是理论数字，实际上PCIe的速度只有理论的一半。PCIe通道通常具有纳秒级别的延迟，因此可以忽略延迟。在总用时上，从4到16个PCIe通道，性能提升约3.2%。但是，如果PyTorch的数据加载器有固定内存，则性能提升为0%。因此，如果使用单个GPU，请不要在PCIe通道上浪费金钱。在选择CPU PCIe通道和主板PCIe通道时，要保证你选择的组合能支持你想要的GPU数量。如果买了支持2个GPU的主板，而且希望用上2个GPU，就要买支持2个GPU的CPU，但不一定要查看PCIe通道数量。

PCIe通道和多GPU并行计算如果在多个GPU上训练网络，PCIe通道是否重要呢？ Tim曾在ICLR 2016上发表了一篇论文指出（<https://arxiv.org/abs/1511.04561>）：如果你有96个GPU，那么PCIe通道非常重要。但是，如果GPU数量少于4个，则根本不必关心PCIe通道。几乎很少有人同时运行超过4个GPU，所以不要在PCIe通道上花冤枉钱。这不重要！

CPU核心数

为了选择CPU，首先需要了解CPU与深度学习的关系。CPU在深度学习中起到什么作用？当在GPU上运行深度网络时，CPU几乎不会进行任何计算。CPU的主要作用有：（1）启动GPU函数调用（2）执行CPU函数。到目前为止，CPU最有用的应用是数据预处理。有两种不同的通用数据处理策略，具有不同的CPU需求。第一种策略是在训练时进行预处理，第二种是在训练之前进行预处理。对于第一种策略，高性能的多核CPU能显著提高效率。建议每个GPU至少有4个线程，即为每个GPU分配两个CPU核心。Tim预计，每为GPU增加一个核心，应该获得大约0-5%的额外性能提升。对于第二种策略，不需要非常好的CPU。建议每个GPU至少有2个线程，即为每个GPU分配一个CPU核心。用这种策略，更多内核也不会让性能显著提升。在图像处理中，我觉得一般是采用第二种！

硬盘/固态硬盘（SSD）

通常，硬盘不会限制深度学习任务的运行，但如果小看了硬盘的作用，可能会让你追、悔、莫、及。想象一下，如果你从硬盘中读取的数据的速度只有100MB/s，那么加载一个32张ImageNet图片构成的mini-batch，将耗时185毫秒。相反，如果在使用数据前异步获取数据，将在185毫秒内加载这些mini-batch的数据，而ImageNet上大多数神经网络的计算时间约为200毫秒。所以，在计算状态时加载下一个mini-batch，性能将不会有任何损失。Tim小哥推荐的是固态硬盘（SSD），他认为SSD在手，舒适度和效率皆有。和普通硬盘相比，SSD程序启动和响应速度更快，大文件的预处理更是要快得多。顶配的体验就是NVMe SSD了，比一般SSD更流畅。

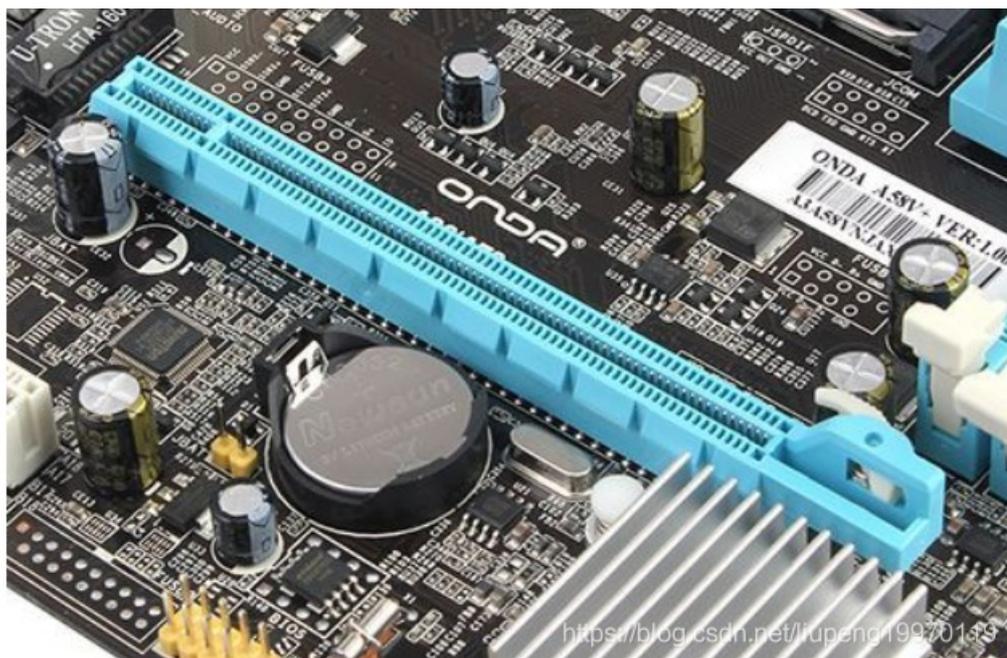
电源装置（PSU）：一个程序员对电源最基础的期望，首先得是能满足各种GPU所需能量吧。随着GPU朝着更低能耗发展，一个质量优秀的PSU能陪你走很久。应该怎么选？Tim小哥有一套计算方法：将电脑CPU和GPU的功率相加，再额外加上10%的功率算作其他组件的耗能，就得到了功率的峰值。举个例子，如果你有4个GPU，每个功率为250瓦，还有一个功率为150瓦的CPU，则需电源提供 $4 \times 250 + 150 + 100 = 1250$ 瓦的电量。

CPU和GPU的冷却

对于产热大户CPU和GPU来说，散热性不好会降低它们的性能。对CPU来说，则标配的散热器，或者AIO水冷散热器都是不错的选择。但GPU该用哪种散热方案，却是个复杂的事。风冷散热如果只有单个GPU，风冷是安全可靠的，但若你GPU多达3-4个，靠空气冷却可能就不能满足需求了。目前的GPU会在运行算法时将速度提升到最大，所以功耗也达到最大值，一旦温度超过80℃，很有可能降低速度，无法实现最佳的性能。建议GPU用水冷。

主板

主板应该有足够的PCIe插槽来支持所需的GPU数量。但需要注意的是，大多数显卡宽度需要占用两个PCIe插槽。



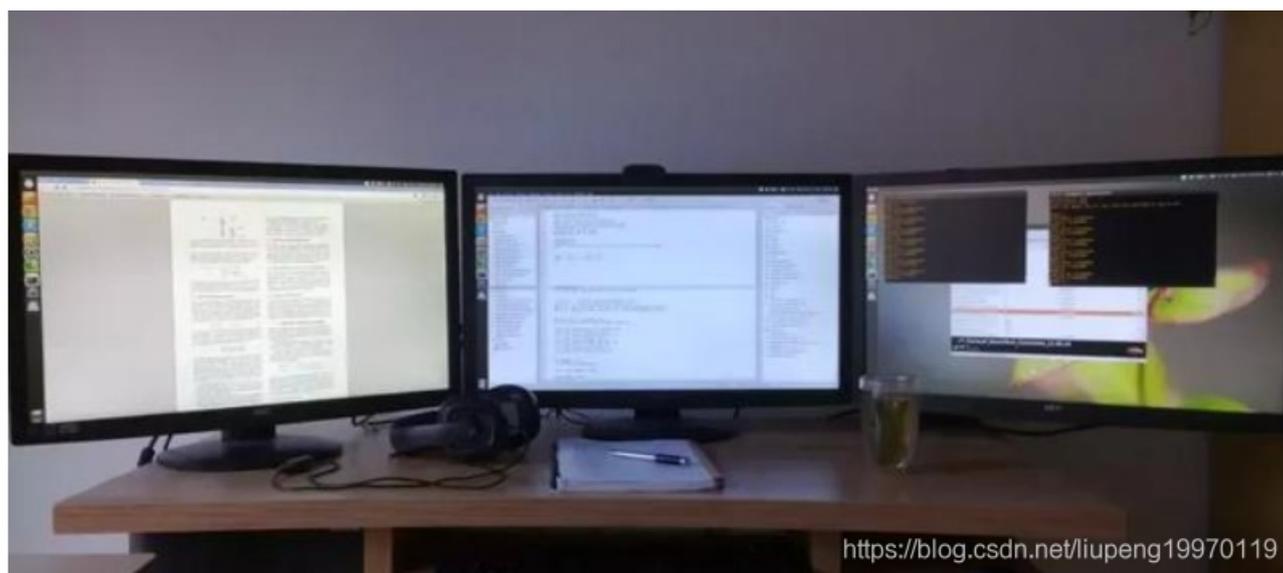
如果打算使用多个GPU，就需要购买PCIe插槽之间有足够空间的主板，确保显卡之间不会相互遮挡。

机箱

选机箱的时候，必须保证机箱能装下主板顶部的全长GPU，虽然大部分机箱是没问题的，但是万一你买小了，那就得看商家给不给你七天无理由了.....所以，买之前最好确认一下机箱的尺寸规格，也可以搜一下机箱装着GPU的图，有别人的成品图的话就能买得放心一些。另外，如果你想用定制水冷的话，保证你的机箱能装得下散热器，尤其是给GPU用定制水冷的时候，每个GPU的散热器都需要占空间。

显示器

三个屏幕就完事了，哈哈。



总结

GPU: RTX 2070、RTX 2080 Ti、GTX 1070、GTX 1080、GTX 1080, 这些都不错。

CPU: 1.给每个GPU配1-2个CPU核心, 具体要看你预处理数据的方式; 2.频率要大于2GHz, CPU要能支持你的GPU数量;
3.PCie通道不重要。

内存: 1.时钟频率无关紧要, 内存买最便宜的就行了; 2.内存 \geq 显存最大的那块GPU的RAM; 3.内存不用太大, 用多少买多少;
4.如果你经常用大型数据集, 买更多内存会很有用。

硬盘/SSD: 1.给你的数据集准备足够大的硬盘 ($\geq 3TB$); 2.有SSD会用得舒坦一些, 还能预处理小型数据集。

PSU: 1.需要的功率最大值 \approx (CPU功率+GPU功率) $\times 110\%$; 2.买一个高效等级的电源, 特别是当你需要连很多GPU并且可能运行很长时间的时, 这样可以节省很多电费; 3.买之前请确保电源上有足够多的接头 (PCie 8-pin或6-pin) 接GPU。

散热: CPU: 标准配置的CPU散热器或者AIO水冷散热器; GPU: 1.单个GPU, 风冷散热即可; 2.若用多个GPU, 选择鼓风式风冷散热或水冷散热。

主板: 准备尽可能多的连接GPU的PCie插槽, 一个GPU需要两个插槽, 每个系统最多4个GPU, 不过你也要考虑GPU的厚度。

显示器: 为了提高效率, 多买几块屏幕吧。

装机是一个一次性学技能的事情, 不要怕麻烦, 学会了以后可以给别人装。最后补充一张显卡建议图:

终极建议

性能最好的GPU: RTX 2080 Ti

性价比高, 但小贵: RTX 2080, GTX 1080

性价比高, 同时便宜: GTX 1070, GTX 1070 Ti, GTX 1060

使用的数据集 $> 250GB$: RTX 2080 Ti, RTX 2080

预算很少: GTX 1060 (6GB)

几乎没预算: GTX 1050 Ti (4GB)/ CPU (建模) + AWS/TPU (训练)

参加Kaggle竞赛: GTX 1060 (6GB) (建模) + AWS (最终训练) + fast ai 库

有前途的CV研究员: GTX 2080 Ti; 在2019年升级到RTX Titan

普通研究员: RTX 2080 Ti/GTX 10XX -> RTX Titan, 注意内存是否合适

有雄心壮志的深度学习菜鸟: 从GTX 1060 (6GB)、GTX 1070、GTX 1070 Ti开始,

随便玩玩的深度学习菜鸟: GTX 1050 Ti (4或2GB)

知乎 @

<https://blog.csdn.net/iupeng19970119>

<https://blog.csdn.net/iupeng19970119>

最终结果

I7 9700K 8核16线程。16G内存，480G固态+2000G机械硬盘。华硕Z390大板。先马600W电源，酷冷至尊大功率风扇，先马透明机箱，罗技办公键鼠，6000整。显卡

花了小1w配了一台1080ti主机。1080ti是技嘉的，3400。