以一半的价格组装专业的深度学习工作站

2019年4月16日更新：可以[在这个帖子](https://l7.curtisnorthcutt.com/the-best-4-gpu-deep-learning-rig)找到一个更好的配置。

2019年3月21日更新：亚马逊的链接添加到所有零部件。新增鼓风机式GPU，更快/更便宜的SSD M.2，和其他选项。

视频地址: https://youtu.be/jLDOibBz5ro

我在麻省理工学院的[量子计算实验室](http://web.mit.edu/~cua/www/quanta/)和[数字学习实验室](https://openlearning.mit.edu/value-digital-learning/research)为研究人员组装了一台多GPU深度学习工作站。当在网上了解如何组装时，找不到足够详细的资料来指导如何选择配件。

我发现可以从网上购买Lambda GPU 工作站之类的整机。唯一的问题：它们的价格为$ 12,500。对于顶级前沿的深度学习研究来说，这是一台很棒的机器，但是如果您预算有限，那就有点为难。这就是为什么我要用[6200美元](https://l7.curtisnorthcutt.com/assets/images/posts/2019/gpu_rig/receipt_gpu_rig.pdf)组装自己的版本的原因 (具有相似或更好的配件) 。为了帮助其他研究人员，我将共享每个配件的详细信息，以便您也可以组装一台。

即使您正在组装较小的深度学习机器或第一台计算机，您也可能会发现本文中的配件说明很有用。在整个过程中，我都提供了一些可以进一步降低成本的例子。

最后，我提供了该机器相对于[Google Compute Engine（GCE）深度学习虚拟机(VM)](https://cloud.google.com/deep-learning-vm/)进行了时间和成本对比。我在[ImageNet](https://github.com/pytorch/examples/tree/master/imagenet)使用[PyTorch的ResNet50训练](https://github.com/pytorch/examples/tree/master/imagenet)进行基准测试。

# 完美的组装？

没有完美的组装，因为每个人的需求都不同。即使有，随着新硬件的发布，完美的组装也在不断变化。这篇文章提供了一个不错的配置，如果在阅读本文时发现有需要改进的，请在下面评论！

# 配件清单

这是[我的配置单](https://l7.curtisnorthcutt.com/assets/images/posts/2019/gpu_rig/receipt_gpu_rig.pdf)。

我通过[Newegg](https://newegg.com/)在线订购了了所有产品，你也可以其他渠道购买（例如[Amazon](https://amzn.to/2UMPVaS)）。如果您附近有[MicroCenter](https://www.microcenter.com/)实体店，那么在实体商店购买时，它们的CPU价格通常很便宜。

深度学习工作站的所有配件。

以下是每个配件以及我支付的价格（2019年1月31日）：

## 3个EVGA RTX 2080 Ti GPU

 [EVGA GeForce RTX 2080 Ti](https://amzn.to/2CuS83K)，   $ 3570（每个$ 1190）  
**更新**：为避免过热，我现在使用以下鼓风机式GPU：  
1. [ASUS GeForce RTX 2080 Ti 11G Turbo Edition GD](https://amzn.to/2uol4pm)，   $ 1209（03/21/2019）  
2. [技嘉GV -N208TTURBO-11GC RTX 2080 Ti Turbo 11GB](https://amzn.to/2Yc0WVc)，   $ 1399（03/21/2019）  
3. [ZOTAC游戏GeForce RTX 2080 Ti Blower 11GB](https://amzn.to/2UMiOUx)，   $ 1299（03/21/2019）

## 20线程CPU（处理器）

 [英特尔酷睿i9-9820X Skylake X 10核3.3Ghz](https://amzn.to/2OhYdoK)，   850美元

## X299主板（所有其他配件与此连接）

 [华硕WS X299 SAGE LGA 2066 Intel X299](https://amzn.to/2Oi2iJz)，   $ 492.26

## 机箱（盒子里的东西都放进去了）

 [Corsair Carbide Series Air 540 ATX机箱](https://amzn.to/2CxFKQn)，   $ 130  
其他选项：[CORSAIR CARBIDE AIR 740 ATX机箱](https://amzn.to/2CxFKQn)，   $ 149（03/21/2019）

## 2TB m.2 SSD（固态硬盘(SSD) - 用于快速访问的持久性存储器）

 [英特尔660p系列M.2 2280 2TB PCI-Express](https://amzn.to/2YegGXZ)，   280美元  
更新：在最近的版本中，我使用以下选项（速度提高2倍，存储减半，价格相近）：  
1. [HP EX950 M.2 1TB PCIe NVMe NAND SSD](https://amzn.to/2CvZxQe)，   199美元（03/21/2019）  
2. [HP EX920 M.2 1TB PCIe NVMe NAND SSD](https://amzn.to/2UNydE7)，   $ 168（03/21/2019）  
3. [三星970 EVO 1TB-NVMe PCIe M.2 SSD](https://amzn.to/2Ok8Gjo)，   $ 250（03/21/2019）

## 3TB硬盘驱动器（机械硬盘 -用于慢速访问的持久性存储）

 [希捷BarraCuda ST3000DM008 3TB 7200 RPM](https://amzn.to/2UNyOFR)，   85美元

## 128GB RAM（内存条）

 [8支CORSAIR Vengeance LPX 16GB](https://amzn.to/2FelPX6)，   $ 740（每支$ 185）

## 1300 1600W PSU（电源）

季节性PRIME 1300W 80+金牌电源，   
 [EVGA SuperNOVA 1600W P2](https://amzn.to/2U35WMS)，   $ 347  
**更新**：正在使用（更便宜）[Rosewill HERCULES 1600W黄金PSU](https://amzn.to/2UN5Vd5)，   $ 209（03/21/19）  
**警告**：在分布式ImageNet / ResNet50基准测试中, 1300W PSU会导致重启动。[其他人也遇到过这个问题。](https://github.com/pytorch/pytorch/issues/3022)增加到1600W后，该问题就解决了。

## 制冷（防止CPU过热）

 [海盗船Hydro Series H100i PRO静音版](https://amzn.to/2FjtNyp)，   110美元

通过Newegg的[企业帐户](https://www.neweggbusiness.com/)享受280美元的折扣并免税后，所有配件的总成本为6200美元（PSU升级费用为+107美元）。



完整的深度学习工作站。

# 配件注意事项

选择GPU，RAM，CPU和主板等配件时，要牢记三个目标：

1. 最大化速度和存储空间。
2. 避免配件瓶颈。
3. 减少开支。

下面，我列出了组装中的每个配件以及每个配件的注意事项。列出配件的顺序是它们对训练深度学习模型的性能影响。

## 显卡

* RTX 2080 Ti在[评测](https://www.videocardbenchmark.net/high_end_gpus.html)中被评为在$ 2500以下的最好GPU
* 购买第三方GPU（例如EVGA或MSI），而不是[Nvidia标准卡](https://www.nvidia.com/en-us/geforce/graphics-cards/rtx-2080-ti/)
* 使用RTX 2080 Ti，请注意[过热问题](https://www.tomshardware.com/news/rtx-2080-ti-gpu-defects-launch,37995.html)。
* 我的组装不使用鼓风机风扇GPU（更便宜），但是鼓风机样式可能会产生[更好的性能](https://www.youtube.com/watch?v=0domMRFG1Rw)。

GPU是任何深度学习机器中最重要的配件。它也是最贵的。应该首先确定要使用哪个GPU：配置中的所有其他内容都将取决于此。有一些[很棒的博客文章，](http://timdettmers.com/2018/12/16/deep-learning-hardware-guide/)介绍了如何根据需要选择合适的GPU。

假设您需要高性能的GPU，我建议您别理营销噱头，只需购买[RTX 2080 Ti即可](https://lambdalabs.com/blog/2080-ti-deep-learning-benchmarks/)。如果您想进行自己的研究并选择一款[物有所值的](https://www.videocardbenchmark.net/high_end_gpus.html) GPU，请查看[videocardbenchmark.net](https://www.videocardbenchmark.net/high_end_gpus.html)并选择价格范围内性能最高的GPU。除非您想支付2500美元以上的价格，否则RTX 2080 Ti是显而易见的选择。如果要降低30％的性能，您可以购买更便宜的RTX 2080或更早的GTX 1080 Ti。无论选择哪种GPU，我都建议购买具有至少11GB内存的GPU，以进行最新的深度学习。这是RTX 2080 Ti的内存量。

购买RTX 2080 Ti时，您会发现有成千上万的品牌：EVGA，技嘉，华硕，MSI…这些被称为第三方品牌GPU。您也可以直接从Nvidia 购买[Nvidia 标准卡](https://www.nvidia.com/en-us/geforce/graphics-cards/rtx-2080-ti/)。通常，如果要获得最佳性能，请**避免**购买标准卡。像EVGA这样的公司会为GPU创建定制案例，以优化性能，有时还会使GPU超频。标准卡是最早的设计，而并非总是最好的尝试。第三方GPU通常具有1-3个风扇，据推测，更多的风扇可以提高性能。其中有一些只是营销噱头，一般有两个风扇就可以。这里的主要信息是：购买EVGA，技嘉，华硕或MSI等第三方品牌GPU。

请注意，在第三方品牌的可选项中，您可能会看到不同的价格。超频的GPU往往会花费更多，但通常会有其他牺牲，实际上并不能带来更好的性能。所以通常买便宜的即可。

一些用户抱怨RTX 2080 TI的[过热问题](https://www.tomshardware.com/news/rtx-2080-ti-gpu-defects-launch,37995.html)。我在组装中仅包含三个GPU的原因是为了增加散热量。如果没有问题，我将添加第四个RTX 2080 TI GPU。

在这个版本我用开放式风扇的GPU，只是因为他们成本低。鼓风机式的GPU直接把热气从机箱侧面排出, 可能会有更好的性能。因为主板的原因，GPU紧紧挨在一起，挡住了开放式GPU风扇。如果您购买鼓风机式GPU，风扇可以将气体从机箱侧面排出。您可以从[此视频](https://www.youtube.com/watch?v=0domMRFG1Rw)了解更多有关这两种风扇差别的信息。

## SSD（固态驱动器）

* SSD <> GPU间的数据传输可以说是深学习训练和预测的主要瓶颈。(译注:考虑到计算时间不少, 并且计算和数据载入是按批量异步的, 实际上SSD速度的重要性没有这么严重, 当然, 低价格的NVMe以及m.2接口都是值得的)
* m.2 SSD可以比标准SSD快7倍。(译注: 要m.2的NVMe协议的才行, m.2是接口, 快是快在NVMe比AHCI快, 而且要Sumsang和Intel之类的读写读写速度才能达到这个水平, 购买时请注意查看读写速度参数, 也要注意甄别有些厂家提供了虚假数据)
* 如果负担得起，请购买m.2 SSD。你需要一个M.2兼容的主板。

从磁盘数据传输到你的GPU是深度学习的主要瓶颈，使用M.2 SSD可以大大减少训练和测试时间。M.2 SSD就是解决方案。最贵的SSD具有3500 MB / s的写入速度，而标准的SSD 是500 MB /秒。

在我的组装中，我购买了便宜的m.2 SSD，其写入速度约为1800 mb / s，但具有2TB的高容量。您会发现买速度快的256G的M.2 SSD也可以。这是以更低的成本获得更好性能的好方法。唯一需要注意的是，要确保M.2 SSD能放得下所有的训练数据。

## 主板

* 为了支持多GPU，你需要足够的PCI-E通道。
* 这意味着你需要一个**x299（英特尔CPU）**或x399（对于AMD CPU）的主板。
* 你可以考虑更便宜的方案，但如果能负担得起，考虑工作站主板。

主板的购买可能会很棘手，因为有很多选择，而且不清楚为什么某些主板的价格要比其他主板高得多。**对于深度学习，主板的最重要的方面是支持PCI-E通道的数量**。在我的配置里，我的主板有44条PCI-E通道。这意味着对于三个GPU的方案（每个都需要16个通道），可以在2个GPU上用16个通道, 在第三个GPU上用8个通道（用掉44个通道中的40个）。大多数的基准测试显示, 如果用8通道替代16通道运行，GPU的性能差异微不足道，但是也许将来会更大的差异。至少确保您的主板能为每个GPU提供所需的最低量的PCI-E通道。因此，对于3 个RTX 2080ti GPU，主板至少需要24条PCI-E通道。

另一个考虑: 要x299（对于英特尔CPU）还是x399（对于AMD的CPU）。英特尔的CPU线程速度较快，但AMD的CPU往往比英特尔的便宜 (就每个线程的价格而言) 。我选择既要速度也要数量（20线程和处理速度快）的Intel处理器，配一个x299主板。

更可靠（更昂贵）的主板通常被称为工作站主板。可靠性的提高是否值得付出代价尚有争议。我选择在我的组装中使用工作站主板，但是如果您想购买更便宜的主板，请查看[SUPERMICRO x299主板](https://www.neweggbusiness.com/product/product.aspx?item=9b-13-183-643)，该[主板](https://www.neweggbusiness.com/product/product.aspx?item=9b-13-183-643)可以满足我的组装的所有需求，但价格要低100美元。

## CPU

* 选择英特尔的X系列（x299主板）或AMD ThreadRipper（x399）
* Intel CPU的每个线程速度更快，但是AMD CPU的每线程价格更低。

通过考虑以下两个问题，根据您的计算需求选择一个CPU：

1. 您是否运行许多多线程作业？
2. 您是否需要每个线程都快？

如果（1）为“是”，但（2）为“否”，则可以使用32线程[AMD Ryzen Threadripper 2950X](https://www.cpubenchmark.net/cpu.php?cpu=AMD+Ryzen+Threadripper+2950X)降低成本。如果对（2）的回答为“是”，则您可能需要Intel CPU。

对于Intel CPU，您需要用于多GPU深度学习的核心Intel X系列CPU。x299主板只能使用X系列CPU，并且您需要x299主板具有足够的PCI-E通道来支持多个GPU。如果仅使用两个GPU，则可以使用便宜的300系列Intel CPU和LGA 1151主板（而不是x299）来降低主板+ CPU的成本。这将在16个PCI通道上运行一个GPU，在8个通道上运行另一个GPU（大多数LGA 1151主板具有24个PCI-E通道，这一点需要再确认下）。

## 机箱

* 选择适合主板的机箱（标准ATX，较小的mini-ATX）。
* 选择一个有空间的机箱以保持GPU凉爽。
* Carbide Series™Air 540高气流ATX立方体机箱是进行深度学习的安全选择。

对于多GPU工作站，气流和冷却就是一切。选择适合您主板的机箱。大多数带有多个GPU的主板都可能是ATX，因此您需要一个适合ATX主板的机箱。如果您不确定要购买哪种机箱，Carbide Series™Air 540高气流ATX立方体机箱是一个不错的选择。

## 机械硬盘

* 如果您的m.2 SSD无法满足存储需求，请购买7200 RPM旋转磁盘。

如果m.2 SSD太小而无法满足持久存储需求，则需要购买机械硬盘。它们相当便宜，速度为5400 RPM（较慢）和7200 RPM（较快）。RPM代表每分钟旋转多少次，这些磁盘实际上在计算机内部旋转，这意味着它们可能很吵。因为它们的成本很低，所以您可能需要7200 RPM磁盘。

## 内存

* 购买外观小的内存，并确保它适合您的情况。
* 避免使用您从未听说过的品牌。

内存的主要考虑因素是数量，物理大小和延迟。在我的版本中，我有128 GB的RAM，但是您可以根据数据集的大小使用64GB或32GB的以降低成本。一般而言，我建议您购买128 GB的（如果可以负担得起的话），以便可以将整个数据集加载到内存中，可避免在深度学习的每轮训练时卡在硬盘<> 内存瓶颈上。

对于多GPU配置，请确保购买外观小的内存（有较小的外壳）。您可以将外观视为内存条的高度。您会在主板上安装很多东西，有时内存的外壳太大的话会影响其他配件的安装。Corsair Vengeance是一个不错的小巧RAM。请注意，对于台式机内存调的长度，通常需要288针内存条。

如果有空着的内存插槽，请注意查看主板文档。使用正确的内存插槽很重要！主板和主板文档通常会给出,基于内存条数量的插槽使用方案。

## PSU（电源）

* 确保您的PSU可以提供足够的电源。这里有[PSU计算器](https://outervision.com/power-supply-calculator)。
* 每个RTX 2080 Ti大约需要300W的功率。
* 选择全模块化，因为更少的电线=更大的气流。
* 我的1300W PSU会导致在最大负载下重启。1600W则可以正常工作。

您可能会看到黄金 PSU与铂金 PSU。这是指PSU的效能（不是金属），其中铂>金>银>青铜>基本与PSU的效率有关。例如，对于相同的计算量，青铜PSU从墙上插座汲取的电能比白金PSU多。如果需要节约电费（并且环保），请考虑购买铂金或金级PSU。

对于此版本，我最初购买了Seasonal PRIME 1300W PSU，但是当我在进行分布式PyTorch ImageNet / ResNet50训练，最大化了所有GPU时，我遇到了[自动重启](https://github.com/pytorch/pytorch/issues/3022)的问题。我换了EVGA SuperNOVA 1600 P2后，该问题解决了。请注意，通过$ sudo nvidia-smi -pl 180命令可以将GPU功率从250W降低到180W ，能够使用1300W PSU工作。这可以工作，但是我建议使用1600W PSU，因为这会限制GPU速度。

## 制冷系统

* 良好的气流和适当的电缆管理通常足以满足GPU冷却的需求。
* 海盗船h100i是高性能（i9 X系列）CPU冷却的安全选择。
* 如果可能，将机器放在凉爽的空调室内。

从机箱风扇到全系统水冷，有许多冷却选项。通常，如果您的箱子较大且电缆管理正确，则不需要花哨的东西。在我的组装中，CPU没有配备散热器，我使用的Corsair h100i在深度学习设备中是相当标准的。较便宜的选择可能是[Noctua NH-U9S CPU散热风扇](https://noctua.at/en/nh-u9s)。我之所以没有购买该风扇，是因为它很大，我不确定它是否会阻塞我的某些RAM插槽。如果仅使用32 GB RAM，则该散热器可以作为更便宜的选择。

# 和Google Compute Engine对比测试

我将该机器与[Google Compute Engine（GCE）深度学习虚拟机(VM)](https://cloud.google.com/deep-learning-vm/)进行了对比测试。这些是预先配置的VM，据称专门[针对深度学习进行了优化](https://cloud.google.com/deep-learning-vm/docs/)。使用GCE深度学习虚拟机专用的CUDA版本和针对其硬件体系结构优化过的驱动程序。GCE VM没有Nvidia RTX 2080 Ti GPU的选项，因此我在VM中使用Tesla K40s配置。根据基准测试任务，Nvidia RTX 2080 Ti的[性能](https://l7.curtisnorthcutt.com/(https:/versus.com/en/nvidia-geforce-rtx-2080-ti-founders-edition-vs-nvidia-tesla-k40))要比GPU Tesla K40 [好2到4倍](https://l7.curtisnorthcutt.com/(https:/versus.com/en/nvidia-geforce-rtx-2080-ti-founders-edition-vs-nvidia-tesla-k40))。因此，为了公平起见，我用该配置上的**一块** RTX 2080 Ti与GCE VM上的**四块** Tesla K40进行对比。

为了进行基准测试，我使用了[PyTorch的Imagenet分布式示例](https://github.com/pytorch/examples/tree/master/imagenet)。我下载了ImageNet 2012训练和验证集，并在我的机器和GCE深度学习VM上运行了以下代码：

python examples/imagenet/main.py -a resnet18 --lr 0.1 --dist-url 'tcp://127.0.0.1:FREEPORT' --dist-backend 'nccl' --multiprocessing-distributed --world-size 1 --rank 0 "/location/where/I/stored/imagenet/"

## GCE深度学习虚拟机 (VM) 规格

对于我创建的虚拟机，规格为：

* 架构：64位，x86\_64
* K40 GPU数量：8 (译注:这里和其他地方说是4块卡不一致)
* 储存空间：394 GB
* 内存：172 GB
* CPU线程数：24

## ImageNet训练时间基准

比较**一轮**的时间：

* 在该装备上每**1个** RTX 2080 TI：   37.5分钟
* GCE VM上每**四个** Tesla K40 GPU（相当于**两个** K80）：   86.3分钟

这些数据是50轮训练的平均值。运行了相同（以上）的命令。两台计算机上均未运行其他进程。

## GCE每轮的费用

对于我使用的GCE体系结构，在充分认识到这不是最具成本效益的配置的情况下，训练费用为：

* **四个** Tesla K40 GPU 每轮$ 12.77

因此，仅针对GPU训练ImageNet的100轮的费用约为1277美元。对于整个虚拟机，**每小时的**成本约为**21美元**。

## 与Lambda的4-GPU工作站进行比较

我所描述的组装旨在优化成本/性能的权衡。如果您想更精确地和Lambda的4-GPU工作站配置对比，则可以查看Lambda的首席执行官Stephen Balaban(真的非常赞) [在Reddit上](https://www.reddit.com/r/MachineLearning/comments/ayd01o/p_i_built_lambdas_12500_deep_learning_rig_for_6200/ei0euft)分享的改动建议：

* 新增一个  ~~1,349~~1,199美元的鼓风机式GPU
* 加$ 159将3个GPU升级为鼓风机：$ 477 （与Newegg / Amazon的3/13/19相同的价格）
* 添加热插拔驱动器托架：+ $ 50
* 使用1600W Rosewill Hercules PSU：-31美元
* 将CPU从10核升级到12核：189美元
* $ 6,200基础价+ $ 1407调整

经过这些调整，Lambda工作站的总成本约为7607美元：比4-GPU Lambda装置低约4500美元。

# 杂项

我使用的操作系统是Ubuntu Server 18.04 LTS。我正在使用带有TensorFlow（从源代码安装）和PyTorch的[Cuda 10.1](https://developer.nvidia.com/cuda-downloads?target_os=Linux)。当我以最大负载长时间使用所有三个GPU时，由于最顶层GPU发生热节流(热保护)，我注意到性能下降了5％-20％。这可能是由于双风扇GPU制冷性能有问题。我建议使用鼓风机式GPU，以避免热节流。

# 致谢和免责声明

非常感谢[Anish Athayle](https://www.anishathalye.com/)，[Xander Breg](https://www.fordbetterworld.org/programs/30-under-30/fellows/xander-breg)，对[Reddit](https://www.reddit.com/r/MachineLearning/comments/ayd01o/p_i_built_lambdas_12500_deep_learning_rig_for_6200/)的评论以及Newegg客户服务（例如Derek Liu）的指导。

免责声明：您发现错误了吗？我可能搞错了很多地方。请在下面评论，以便我们可以改进！