

R blogs / lang

Articles about R, in your own language

Home

About

Add your blog

Contact us

now starting at

\$40

per line/4lines

Per month.Plus taxes and fees.
When you enroll in Auto Pay.

Shop now

如何搭建一台深度学习

October 17, 2015

is kindly contributed by [数据](#)
[python](#) - go there to comment
[the full post.](#)

早期，一名极客的满足感很大程度上来自于自己搭建的一台机器。到了深度学习的时代，前面那句话仍然是对的。

缘起

在2013年，MIT科技评论将深度学习列为当年十大科技突破之首。其原因在于，模型有其为庞大的网络结构，参数够多，学习能力强，能配合大数据达到惊人的效果。而且，能自动学习特征，避免了“特征工程”这种繁琐的手工劳动。对于图像、音频和文字处理领域有极大的意义。因为最近在尝试用深度学习做文本挖掘，所以需要一台深度学习服务器（相信我，如果用CPU来跑，你的人生显得好短）。

那么就有三个选择项：

A、购买组装好的服务器，例如NVIDIA DIGITS

RECENT POSTS

- [Z](#)
- [p](#)
- [a](#)
- [m](#)
- [i](#)
- [e](#)
- [t](#)
- [n](#)
- [i](#)
- [k](#)
- [a](#)
- [n](#)
- [a](#)
- [u](#)
- [c](#)
- [z](#)
- [y](#)
- [c](#)
- [i](#)
- [e](#)
- [l](#)

LANGUAGES

- [Chinese](#)
- [Dutch](#)
- [French](#)
- [German](#)
- [Indonesian](#)
- [Italian](#)
- [Korean](#)
- [Polish](#)
- [portuguese](#)
- [Russian](#)
- [Serbian](#)

- [Spanish](#)
- [Uncategorized](#)

[DEVBOX](#)

B、购买云服务，例如Amazon的[GPU服务](#)

C、自己攒一台深度学习服务器。

A项从性价比上不合算，而且不一定买的到。B项的云服务对于研究探索性质工作而言也比较贵，机器配置也不一定合适，所以我们就C项了。

硬件选择：基本思路是单显卡机器，保留升级空间

显卡选择：

先解释一点，深度学习为什么需要显卡计算？

GPU 是为大规模的并行运算而优化；

GPU 上则更多的是运算单元（整数、浮点的乘加单元，特殊运算单元等等）；

GPU 往往拥有更大带宽的显存，因此在大吞吐量的应用中也会有很好的性能。

这里有一个很有趣的[解释视频](#)，解释了GPU和CPU的计算区别。

所以显卡选择最重要，因为准备基于CUDA计算（CUDA (Compute Unified Device Architecture) 是NVIDIA开发的GPU并行计算环境。），所以只能选择Nvida系列的。在英伟达产品系列中，有消费领域的GeForce系列，有专业绘图领域的Quadro系列，有高性能计算领域的Tesla系列，如何选择？

有论文研究，太多的精度对于深度学习的错误率是没有提升的，而且大部分的环境框架都只支持单精度，所以双精度浮点计算是不必要，Tesla系列都去掉了。从显卡效能的指标看，CUDA核心数要多，GPU频率要快，显存要大，带宽要

a
a
k
a
d
e
m
i
c
k
i
e
g
o
—
lr
r
a
c
j
o
n
a
l
n
e
w
y
b
o
r
y
• L
,

高。这样，最新Titan X算是价格便宜量又足的选择。

CPU选择：

在深度学习任务中，CPU并不负责主要任务，单显卡计算时只有一个核心达到100%负荷，所以CPU的核心数量和显卡数量一致即可，太多没有必要，但是处理PCIE的带宽要到40。

主板选择：

需要支持X99架构，支持PCIE3.0，还要支持4通道DDR4内存架构。如果要搞四显卡并行，PCIE带宽支持要达到40，并且支持4-Way NVIDA SLI技术。

内存：

达到显存的二倍即可，当然有钱的话越大越好。

电源问题：一个显卡的功率接近300W，四显卡建议电源在1500W以上，为了以后扩展，选择了1600W的电源。

机箱散热：

因为各种部件相当庞大，需要有良好散热功能的大机箱，选择了Tt Thermaltake Core V51机箱，标配3个12cm风扇。未来如果需要还可以加装水冷设备。

以上就是主要硬件环境的选择和安装。

软件环境安装：

主要安装了Ubuntu系统，CUDA环境，以及theano、keras环境

é
c
a
rt
d
,
â
g
e
e
n
tr
e
c
o
n
j
o
i
n
t
s
• 9
t
h
M
il
a
n
o
R
M
e
e

1、安装ubuntu 14.04（不要安装麒麟版本）

通过官网下载iso文件刻录U盘安装，在SSD硬盘上自动安装完成。

2、安装cuda 7.5 deb文件（自带显卡驱动）

2.1 获取CUDA安装包，在NVidia官网下载local deb文件。

2.2执行如下命令安装：

```
$ sudo dpkg -i cuda-repo-ubuntu1410-7.0-  
local_7.0-28_amd64.deb
```

```
$ sudo apt-get update
```

```
$ sudo apt-get install cuda
```

2.3设置环境变量：sudo gedit /etc/profile

```
$ export PATH=/usr/local/cuda-7.5/bin:$PATH
```

```
$ export LD_LIBRARY_PATH=/usr/local  
/cuda-7.5/lib64:$LD_LIBRARY_PATH
```

2.4重启机器：用其中一个样例进行验证测试

```
$ cuda-install-samples-7.5.sh ~
```

```
$ cd ~/NVIDIA_CUDA-Samples_7.5  
/5_Simulations/nbody
```

```
$ make
```

```
$ ./nbody
```

或者在sample目录中make测试项目1_utility，再运行./deviceQuery

3、安装python以及theano工具

3.1 直接下载安装anaconda套件64位版本。

下载后用conda升级到最新。anaconda自带openblas，使numpy和theano计算能达到最大效能

3.2 修改了theano配置文件，使之缺省由gpu运算

在主目录下新建.theanorc文件

.theanorc中输入下面的内容

ti
n
g
:
N
o
v
e
m
b
e
r
2
0
t
h
• O
6
3
o
p
i
H
T
e
p
e
c
H
bl
x
R
-
n

```
[cuda]
root=/usr/local/cuda/bin/
```

```
[global]
floatX = float32
device = gpu0
```

```
[nvcc]
fastmath = True
```

3.3 安装完成后参考如下地址进行测试

```
THEANO_FLAGS=floatX=float32,device=gpu
python `python -c "import os, theano; print
os.path.dirname(theano.__file__)"`/misc
/check_blas.py
```

3.4 安装keras

通过anaconda安装pip install keras

注意keras依赖的theano是dev版本的，需要从github下源码来安装。注意本系统此时有两个python，所以要写路径

```
sudo /home/user/anaconda/bin/python
setup.py develop
```

4、设置远程服务器调用

4.1 安装ssh，启动服务

4.2 设置notebook server

在ipython中如下命令设置密码：

```
from IPython.lib import passwd
passwd()
```

记下生成的字符串。

创建一个ipython配置文件，比如起名叫myserver

```
ipython profile create myserver
vim ~/.ipython/profile_myserver
```

/ipython_notebook_config.py

编辑文件，加入下面几项：

```
c = get_config()
```

a
k
e
t
o
B
3
a
o
k
T
я
б
р
ь
2
0
1
7
г.
• 9
t
h
M
il
a
n
o
R
m
e
e
ti
n

```
c.IPKernelApp.pylab = 'inline' #启动inline模式
c.NotebookApp.ip = '*'
c.NotebookApp.open_browser = False
c.NotebookApp.password =
u'sha1:yourhashedpassword' #把第二步的密码
考进来
c.NotebookApp.port = 6868 #自己设一个端口
号
启动服务
ipython notebook --profile=myserver
```

4.3 远程浏览器调用

远程中浏览器登陆<http://192.168.3.31:6868/> ,
输入密码, 即可进入ipython notebook。
如果需保持连接, 则
nohup ipython notebook --profile=myserver
杀掉连接
ls -l nohup.out
kill -9 "PID"

完工!

最后的硬件配置:

CPU: Intel X99平台 i7 5960K
内存: DDR4 2800 32G(8G*4)
主板: GIGABYTE X99-UD4
显卡: GTX Titan X
硬盘: SSD+普通硬盘

系统和软件

操作系统: Ubuntu 14.04.3 x64
CUDA: 7.5
Anaconda 2.3
Theano 7.0

g
o
n
N
o
v
e
m
b
e
r
2
0
t
h
:
c
a
ll
f
o
r
p
r
e
s
e
n
t
a
ti
o
n
s

Keras 2.0

!

参考资料：

<http://timdettmers.com/2014/08/14/which-gpu-for-deep-learning/>

<http://timdettmers.com/2015/03/09/deep-learning-hardware-guide/>

<http://graphific.github.io/posts/running-a-deep-learning-dream-machine/>

http://docs.nvidia.com/cuda/cuda-quick-start-guide/index.html#ubuntu-x86_64-deb

http://deeplearning.net/software/theano/tutorial/using_gpu.html#cuda

之前网上找到另一种安装theano方式是[自行安装组件](#)。但尝试未成功，有可能是因为CPU等硬件不同，可能需要自行编译安装openblas，numpy，scipy，theano等环境，极为繁琐。最后还是直接用Anaconda方便。

← Web
Programmer in
Milan, Italy

MilanoR meeting:
call for
presentations →

☺