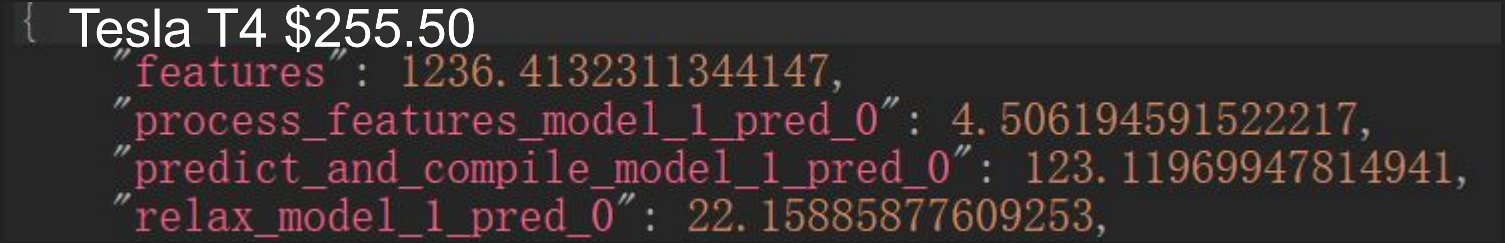
Google Cloud Alphafold Server Statistics

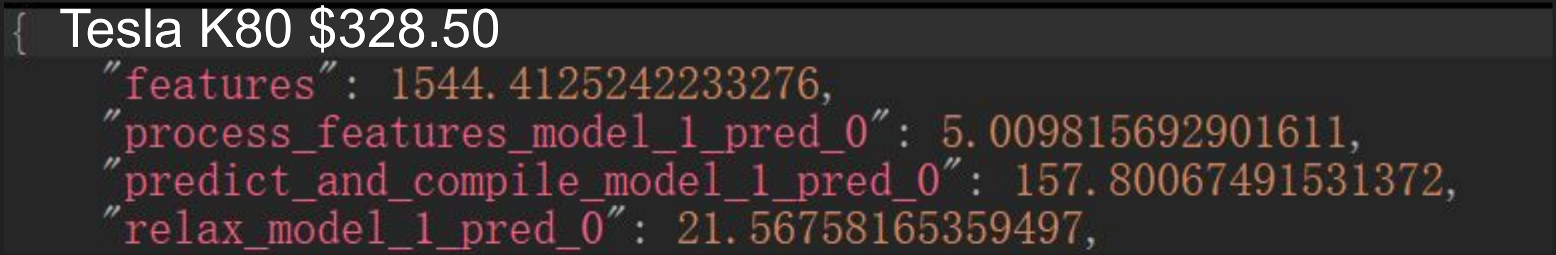
Jason Zhu

Single Process Performance



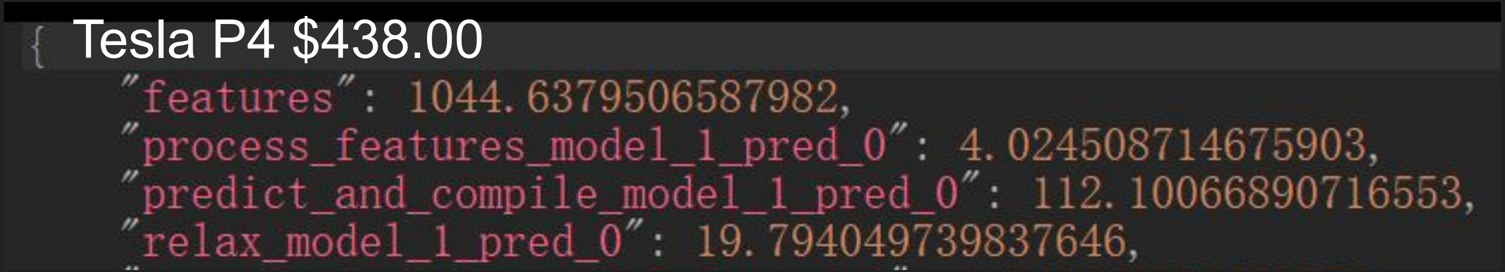
Tesla T4 $255.50

```
{
    "features": 1236.4132311344147,
    "process_features_model_1_pred_0": 4.506194591522217,
    "predict_and_compile_model_1_pred_0": 123.11969947814941,
    "relax_model_1_pred_0": 22.15885877609253,
```



Tesla K80 $328.50

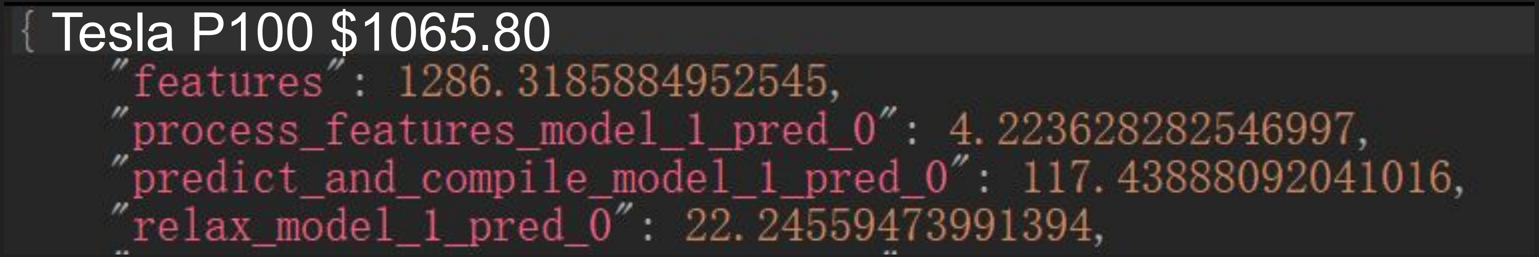
```
{
    "features": 1544.4125242233276,
    "process_features_model_1_pred_0": 5.009815692901611,
    "predict_and_compile_model_1_pred_0": 157.80067491531372,
    "relax_model_1_pred_0": 21.56758165359497,
```



Tesla P4 $438.00

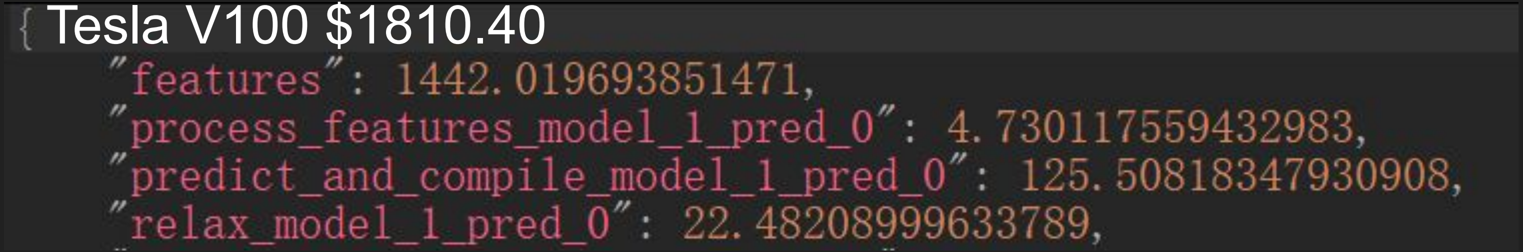
```
{
    "features": 1044.6379506587982,
    "process_features_model_1_pred_0": 4.024508714675903,
    "predict_and_compile_model_1_pred_0": 112.10066890716553,
    "relax_model_1_pred_0": 19.794049739837646,
```

Single Process Performance



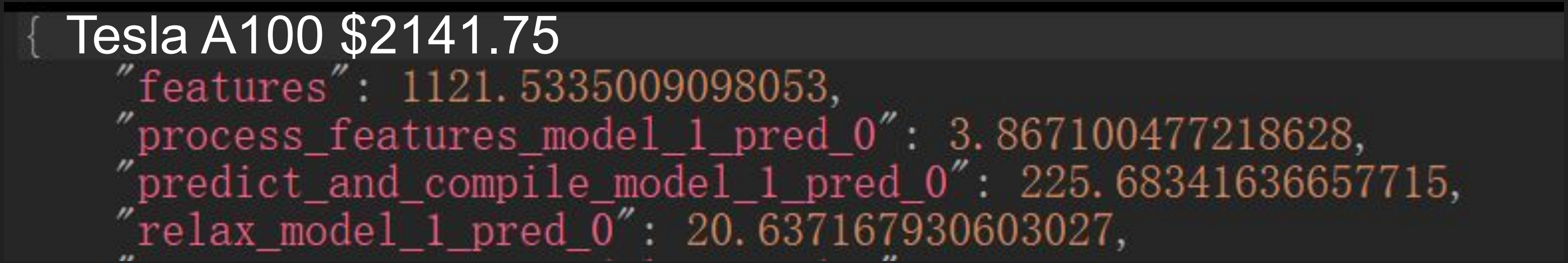
Tesla P100 $1065.80

```
{
    "features": 1286.3185884952545,
    "process_features_model_1_pred_0": 4.223628282546997,
    "predict_and_compile_model_1_pred_0": 117.43888092041016,
    "relax_model_1_pred_0": 22.24559473991394,
```



Tesla V100 $1810.40

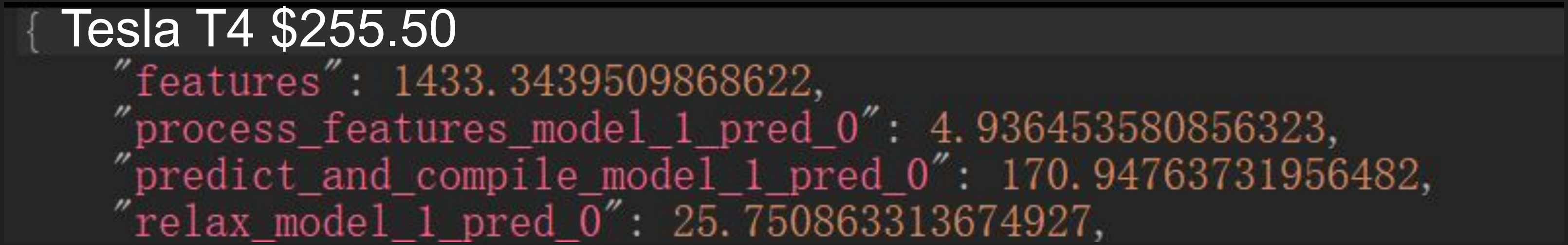
```
{
    "features": 1442.019693851471,
    "process_features_model_1_pred_0": 4.730117559432983,
    "predict_and_compile_model_1_pred_0": 125.50818347930908,
    "relax_model_1_pred_0": 22.48208999633789,
```



Tesla A100 $2141.75

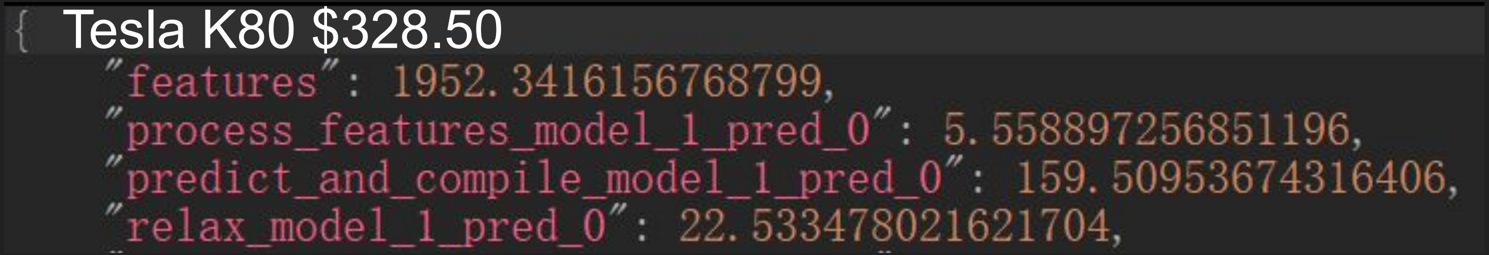
```
{
    "features": 1121.5335009098053,
    "process_features_model_1_pred_0": 3.867100477218628,
    "predict_and_compile_model_1_pred_0": 225.68341636657715,
    "relax_model_1_pred_0": 20.637167930603027,
```

Multiple Process (3 processes) Performance



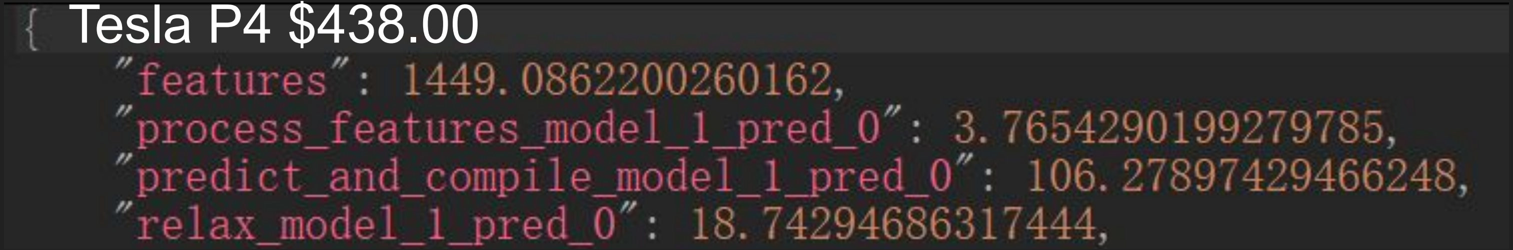
Tesla T4 $255.50

```
{
    "features": 1433.3439509868622,
    "process_features_model_1_pred_0": 4.936453580856323,
    "predict_and_compile_model_1_pred_0": 170.94763731956482,
    "relax_model_1_pred_0": 25.750863313674927,
```



Tesla K80 $328.50

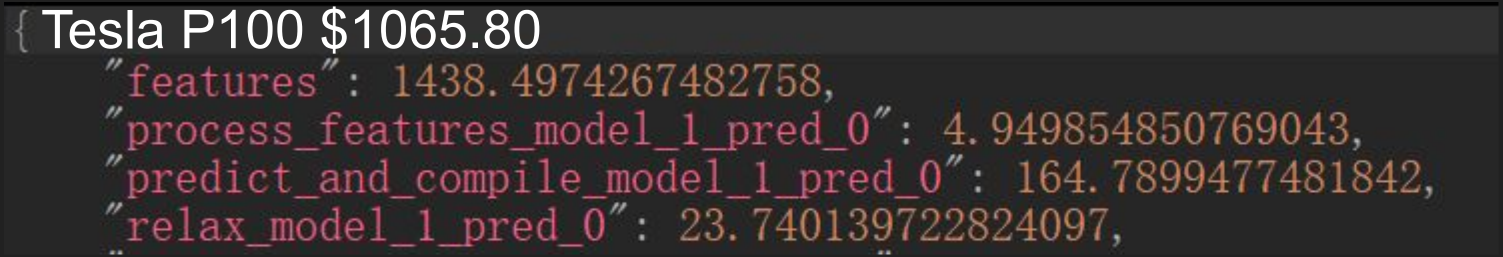
```
{
    "features": 1952.3416156768799,
    "process_features_model_1_pred_0": 5.558897256851196,
    "predict_and_compile_model_1_pred_0": 159.50953674316406,
    "relax_model_1_pred_0": 22.533478021621704,
```



Tesla P4 $438.00

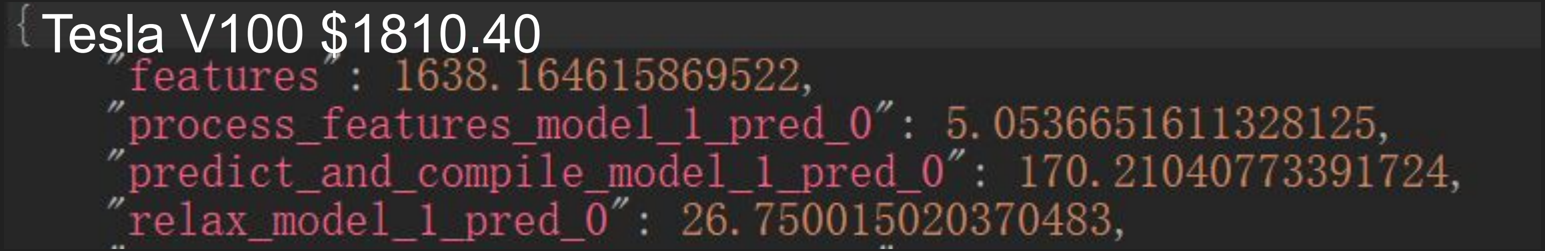
```
{
    "features": 1449.0862200260162,
    "process_features_model_1_pred_0": 3.7654290199279785,
    "predict_and_compile_model_1_pred_0": 106.27897429466248,
    "relax_model_1_pred_0": 18.74294686317444,
```

Multiple Process (3 processes) Performance



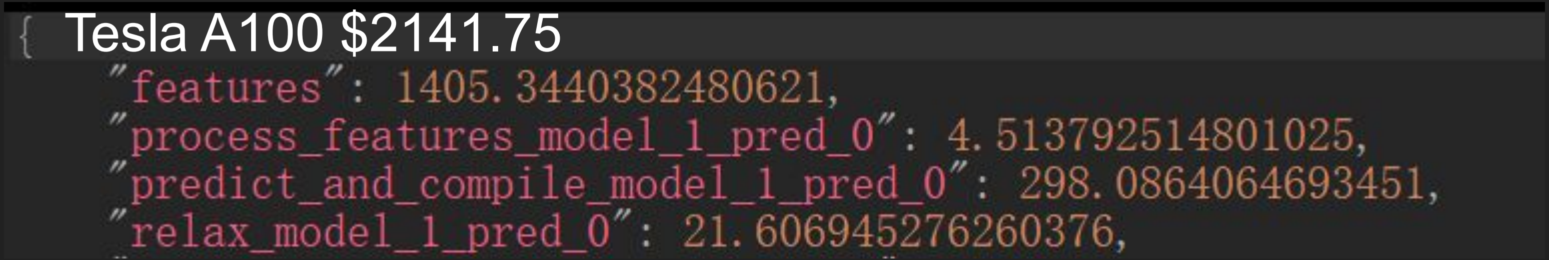
Tesla P100 $1065.80

```
{
    "features": 1438.4974267482758,
    "process_features_model_1_pred_0": 4.949854850769043,
    "predict_and_compile_model_1_pred_0": 164.7899477481842,
    "relax_model_1_pred_0": 23.740139722824097,
```



Tesla V100 $1810.40

```
{
    "features": 1638.164615869522,
    "process_features_model_1_pred_0": 5.0536651611328125,
    "predict_and_compile_model_1_pred_0": 170.21040773391724,
    "relax_model_1_pred_0": 26.750015020370483,
```



Tesla A100 $2141.75

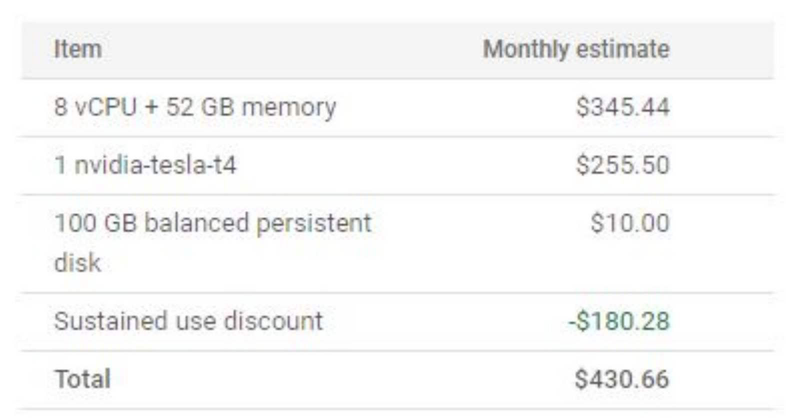
```
{
    "features": 1405.3440382480621,
    "process_features_model_1_pred_0": 4.513792514801025,
    "predict_and_compile_model_1_pred_0": 298.0864064693451,
    "relax_model_1_pred_0": 21.606945276260376,
```

General Analysis

- there’s no clear relationship between gpu prices and running time
- one process takes maximum of 225% vCPU resource each process, so for setting of 8 vCPUs & 52GB RAM, it supports 3 simultaneous processes.
- gpu can help to boost the overall speed, while the effect of multiple gpu is not yet clear
- adding more CPU & correponding RAM would increase the maximum simultaneous process - $33 for each vCPU&RAM on google cloud

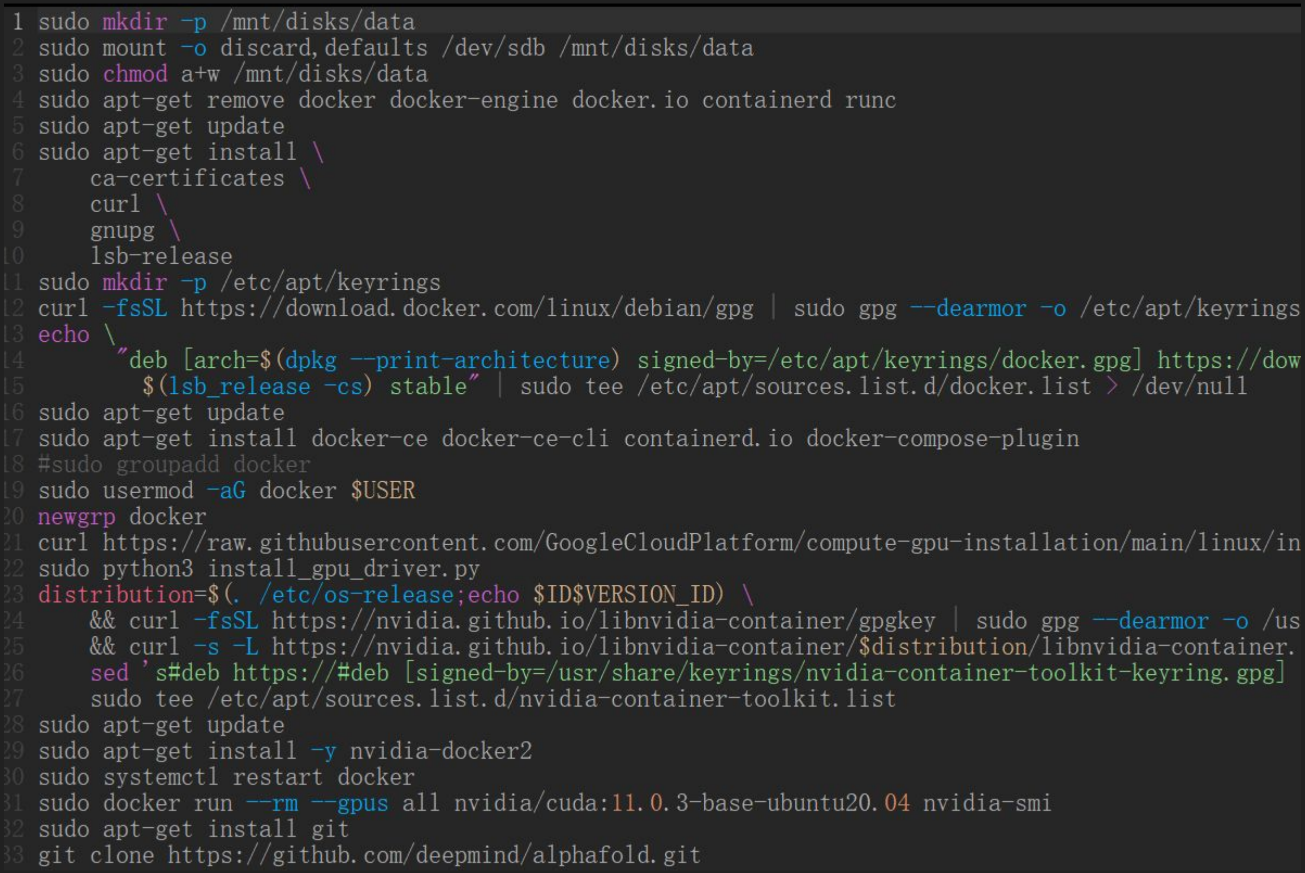
Recommandation Based on Current Data

- 3TB SSD disk for data storage - $307.2 monthly
- 8 vCPU & 52GB RAM ($345.44) with 1 teslaT4 GPU ($255.50) as a basic set
- 100 GB boot disk ($10 monthly)
- 725 seconds for one GFP11 mutation prediction (16 amino acids)
- to increase running speed, only need to add more basic set with no need for SSD and boot disk



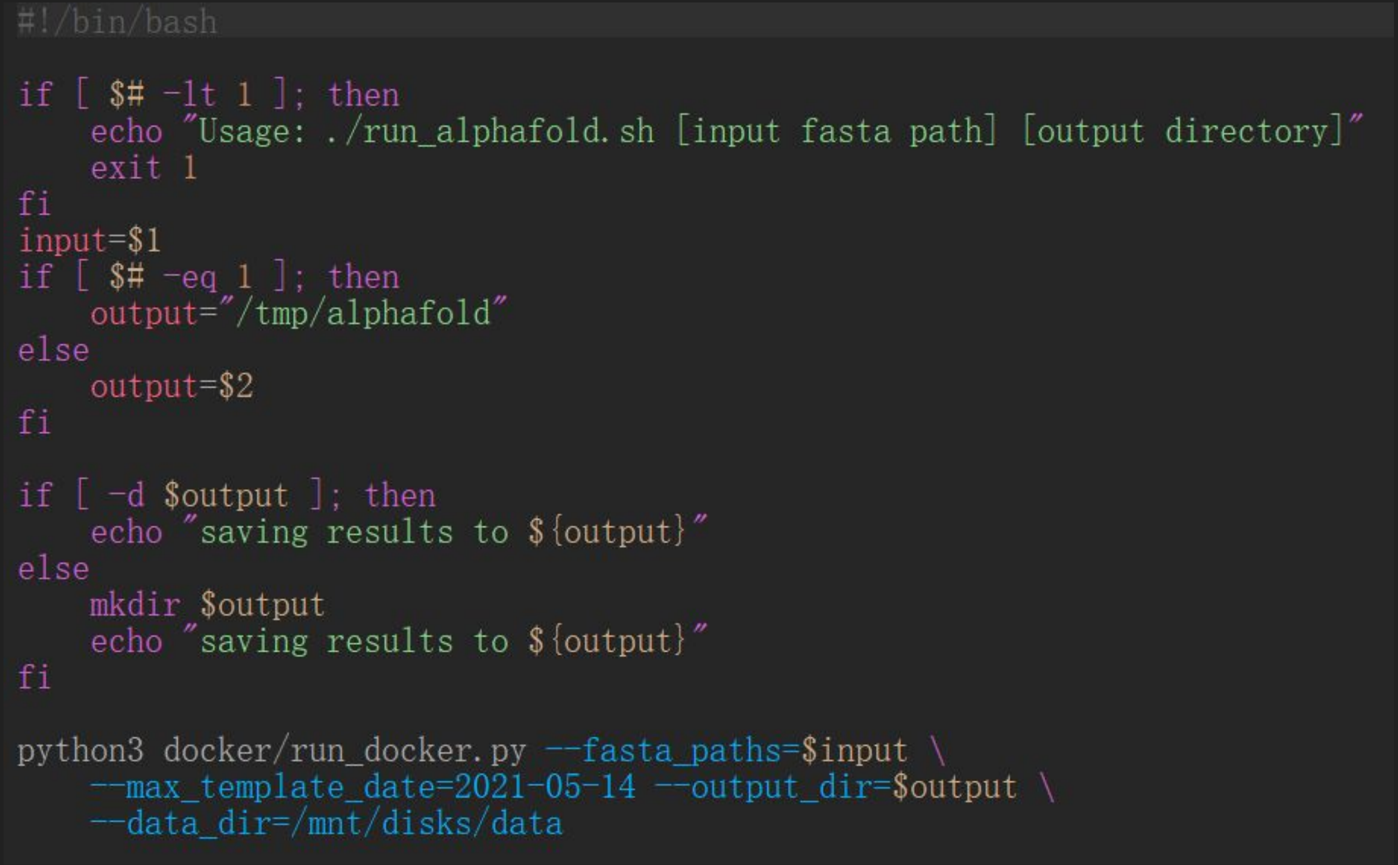
Item	Monthly estimate
8 vCPU + 52 GB memory	$345.44
1 nvidia-tesla-t4	$255.50
100 GB balanced persistent disk	$10.00
Sustained use discount	-$180.28
Total	$430.66

shell script for auto-initialization



```
sudo mkdir -p /mnt/disks/data
sudo mount -o discard,defaults /dev/sdb /mnt/disks/data
sudo chmod a+w /mnt/disks/data
sudo apt-get remove docker docker-engine docker.io containerd runc
sudo apt-get update
sudo apt-get install \
    ca-certificates \
    curl \
    gnupg \
    lsb-release
sudo mkdir -p /etc/apt/keyrings
curl -fsSL https://download.docker.com/linux/debian/gpg | sudo gpg --dearmor -o /etc/apt/keyrings
echo \
      "deb [arch=$(dpkg --print-architecture) signed-by=/etc/apt/keyrings/docker.gpg] https://dow
        $(lsb_release -cs) stable" | sudo tee /etc/apt/sources.list.d/docker.list > /dev/null
sudo apt-get update
sudo apt-get install docker-ce docker-ce-cli containerd.io docker-compose-plugin
#sudo groupadd docker
sudo usermod -aG docker $USER
newgrp docker
curl https://raw.githubusercontent.com/GoogleCloudPlatform/compute-gpu-installation/main/linux/in
sudo python3 install_gpu_driver.py
distribution=$(. /etc/os-release;echo $ID$VERSION_ID) \
    && curl -fsSL https://nvidia.github.io/libnvidia-container/gpgkey | sudo gpg --dearmor -o /us
    && curl -s -L https://nvidia.github.io/libnvidia-container/$distribution/libnvidia-container.
    sed 's#deb https://#deb [signed-by=/usr/share/keyrings/nvidia-container-toolkit-keyring.gpg]
    sudo tee /etc/apt/sources.list.d/nvidia-container-toolkit.list
sudo apt-get update
sudo apt-get install -y nvidia-docker2
sudo systemctl restart docker
sudo docker run --rm --gpus all nvidia/cuda:11.0.3-base-ubuntu20.04 nvidia-smi
sudo apt-get install git
git clone https://github.com/deepmind/alphafold.git
```

Shell script for one step running



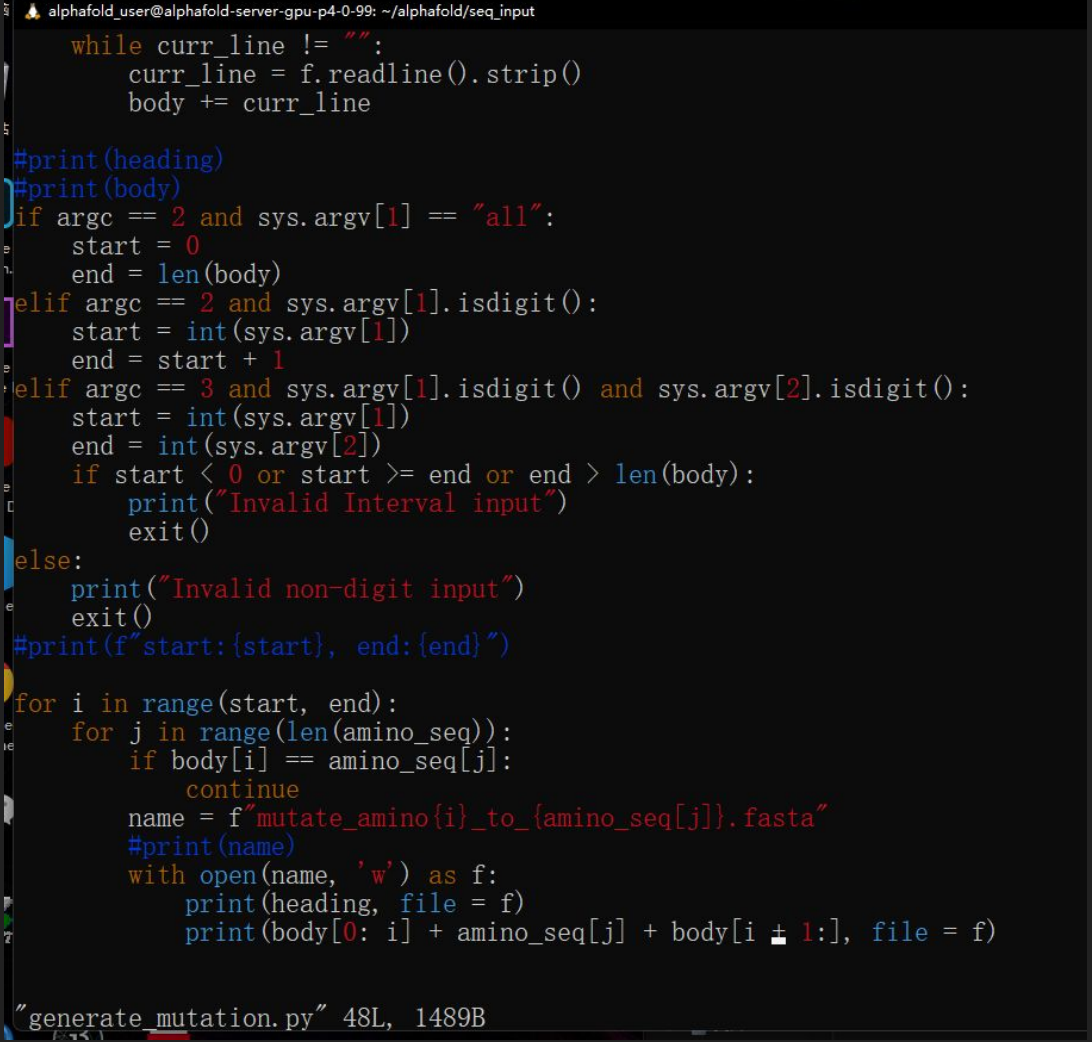
```
#!/bin/bash

if [ $# -lt 1 ]; then
    echo "Usage: ./run_alphafold.sh [input fasta path] [output directory]"
    exit 1
fi
input=$1
if [ $# -eq 1 ]; then
    output="/tmp/alphafold"
else
    output=$2
fi

if [ -d $output ]; then
    echo "saving results to ${output}"
else
    mkdir $output
    echo "saving results to ${output}"
fi

python3 docker/run_docker.py --fasta_paths=$input \
    --max_template_date=2021-05-14 --output_dir=$output \
    --data_dir=/mnt/disks/data
```

Python script for auto-mutation generation



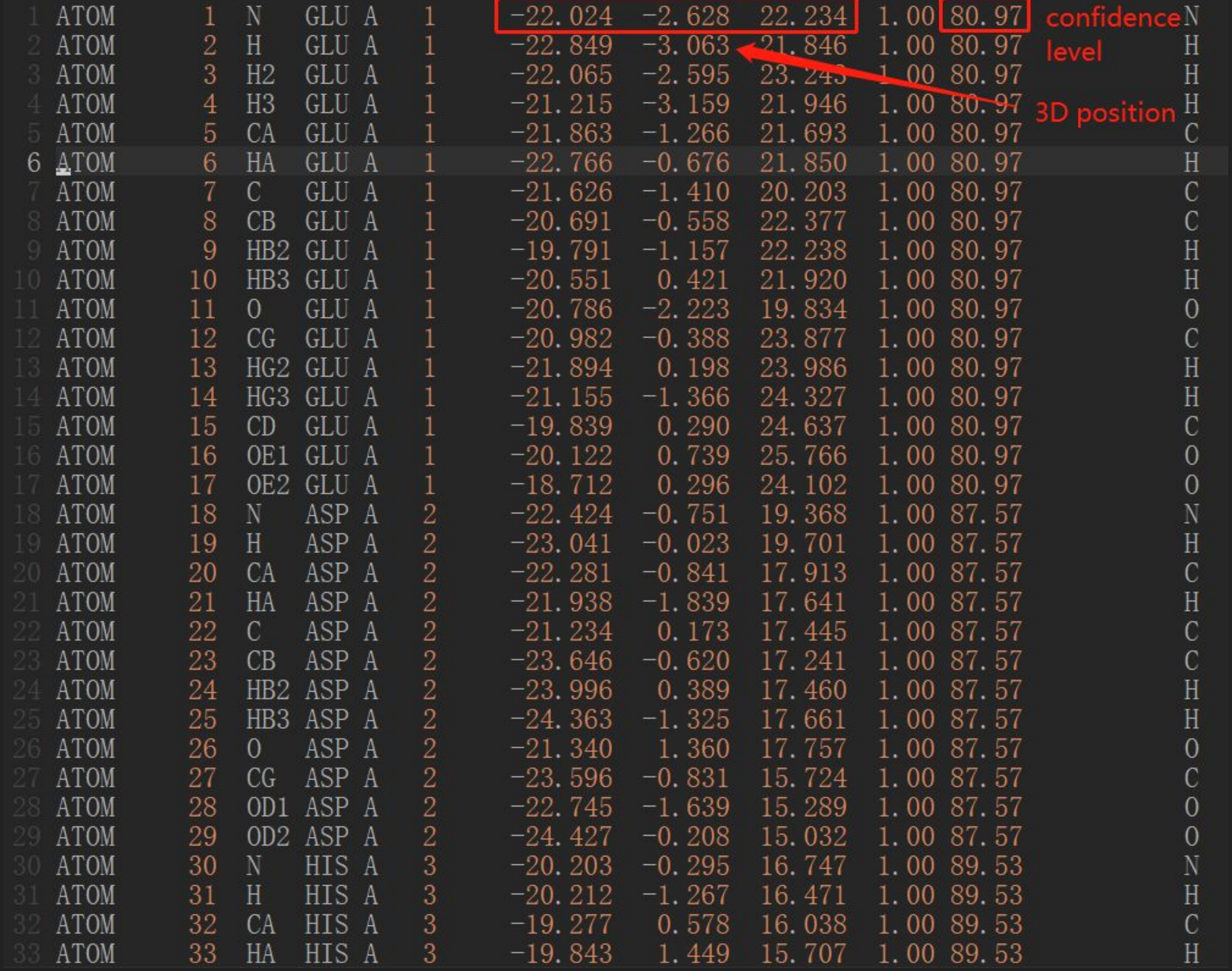
alphafold_user@alphafold-server-gpu-p4-0-99: ~/alphafold/seq_input

```
    while curr_line != "":
        curr_line = f.readline().strip()
        body += curr_line

#print(heading)
#print(body)
if argc == 2 and sys.argv[1] == "all":
    start = 0
    end = len(body)
elif argc == 2 and sys.argv[1].isdigit():
    start = int(sys.argv[1])
    end = start + 1
elif argc == 3 and sys.argv[1].isdigit() and sys.argv[2].isdigit():
    start = int(sys.argv[1])
    end = int(sys.argv[2])
    if start < 0 or start >= end or end > len(body):
        print("Invalid Interval input")
        exit()
else:
    print("Invalid non-digit input")
    exit()
#print(f"start:{start}, end:{end}")

for i in range(start, end):
    for j in range(len(amino_seq)):
        if body[i] == amino_seq[j]:
            continue
        name = f"mutate_amino{i}_to_{amino_seq[j]}.fasta"
        #print(name)
        with open(name, 'w') as f:
            print(heading, file = f)
            print(body[0: i] + amino_seq[j] + body[i + 1:], file = f)

"generate_mutation.py" 48L, 1489B
```



```
ATOM      1  N   GLU A   1     -22.024  -2.628  22.234  1.00 80.97           N
ATOM      2  H   GLU A   1     -22.849  -3.063  21.846  1.00 80.97           H
ATOM      3  H2  GLU A   1     -22.065  -2.595  23.243  1.00 80.97           H
ATOM      4  H3  GLU A   1     -21.215  -3.159  21.946  1.00 80.97           H
ATOM      5  CA  GLU A   1     -21.863  -1.266  21.693  1.00 80.97           C
ATOM      6  HA  GLU A   1     -22.766  -0.676  21.850  1.00 80.97           H
ATOM      7  C   GLU A   1     -21.626  -1.410  20.203  1.00 80.97           C
ATOM      8  CB  GLU A   1     -20.691  -0.558  22.377  1.00 80.97           C
ATOM      9  HB2 GLU A   1     -19.791  -1.157  22.238  1.00 80.97           H
ATOM     10  HB3 GLU A   1     -20.551   0.421  21.920  1.00 80.97           H
ATOM     11  O   GLU A   1     -20.786  -2.223  19.834  1.00 80.97           O
ATOM     12  CG  GLU A   1     -20.982  -0.388  23.877  1.00 80.97           C
ATOM     13  HG2 GLU A   1     -21.894   0.198  23.986  1.00 80.97           H
ATOM     14  HG3 GLU A   1     -21.155  -1.366  24.327  1.00 80.97           H
ATOM     15  CD  GLU A   1     -19.839   0.290  24.637  1.00 80.97           C
ATOM     16  OE1 GLU A   1     -20.122   0.739  25.766  1.00 80.97           O
ATOM     17  OE2 GLU A   1     -18.712   0.296  24.102  1.00 80.97           O
ATOM     18  N   ASP A   2     -22.424  -0.751  19.368  1.00 87.57           N
ATOM     19  H   ASP A   2     -23.041  -0.023  19.701  1.00 87.57           H
ATOM     20  CA  ASP A   2     -22.281  -0.841  17.913  1.00 87.57           C
ATOM     21  HA  ASP A   2     -21.938  -1.839  17.641  1.00 87.57           H
ATOM     22  C   ASP A   2     -21.234   0.173  17.445  1.00 87.57           C
ATOM     23  CB  ASP A   2     -23.646  -0.620  17.241  1.00 87.57           C
ATOM     24  HB2 ASP A   2     -23.996   0.389  17.460  1.00 87.57           H
ATOM     25  HB3 ASP A   2     -24.363  -1.325  17.661  1.00 87.57           H
ATOM     26  O   ASP A   2     -21.340   1.360  17.757  1.00 87.57           O
ATOM     27  CG  ASP A   2     -23.596  -0.831  15.724  1.00 87.57           C
ATOM     28  OD1 ASP A   2     -22.745  -1.639  15.289  1.00 87.57           O
ATOM     29  OD2 ASP A   2     -24.427  -0.208  15.032  1.00 87.57           O
ATOM     30  N   HIS A   3     -20.203  -0.295  16.747  1.00 89.53           N
ATOM     31  H   HIS A   3     -20.212  -1.267  16.471  1.00 89.53           H
ATOM     32  CA  HIS A   3     -19.277   0.578  16.038  1.00 89.53           C
ATOM     33  HA  HIS A   3     -19.843   1.449  15.707  1.00 89.53           H
```

confidence level

3D position

shell/python program for automatic running

- ssh/mobaxterm for login, all have linux interface avaliable
- scp for transferring data back / store in server locally