针对问题：

在十一号机子上，数据读取的时间占了总算法运行时间的一半甚至更多，在提升 dataloader 的 num_works 同时并没有改善。 数据读取占用很久时间的现象理应是不正常的，同时在监控 GPU 运行效率是会发现 GPU 会浪费很多时间在等待读取完的数据输入，即 GPU 利用率会间隔一定时间持续为；同时，这种现象在小规模数据集上不会对总体运行时间带来很大的影响，但在大规模数据集上的每个 batch 运行时间的累加会带来非常大的影响。

怀疑因素：

1. 11 号服务器上的硬盘不是直连的，通过 nas 连接，在每次读取 batch 数据的时候，I/O 口的性能瓶颈；
2. 数据处理的代码段消耗太多时间。

实验：

在 11，7，8 号机子上做了相同的读取数据实验，并分别记录了硬盘读取和数据预处理的时间，并在不同的 num_works(0，1，2，4)的条件下进行实验。实际测算结果如下表所示：

Batchsize：64

服务器	0			1			2			4		
	总数据时间	硬盘读取时间	数据处理时间	总数据时间	硬盘读取时间	数据处理时间	总数据时间	硬盘读取时间	数据处理时间	总数据时间	硬盘读取时间	数据处理时间
11 号机	4.514	0.223	3.048	0.2699	0.0737	0.1419	0.2042	0.0823	0.1535	0.1591	0.0876	0.1662
7 号机	0.2602	0.0532	0.1732	0.3010	0.0550	0.1800	0.2438	0.0617	0.2176	0.2525	0.0654	0.2016
8 号机	0.1196	0.0268	0.0778	0.1493	0.0280	0.0917	0.1345	0.0654	0.1026	0.1212	0.0350	0.0916

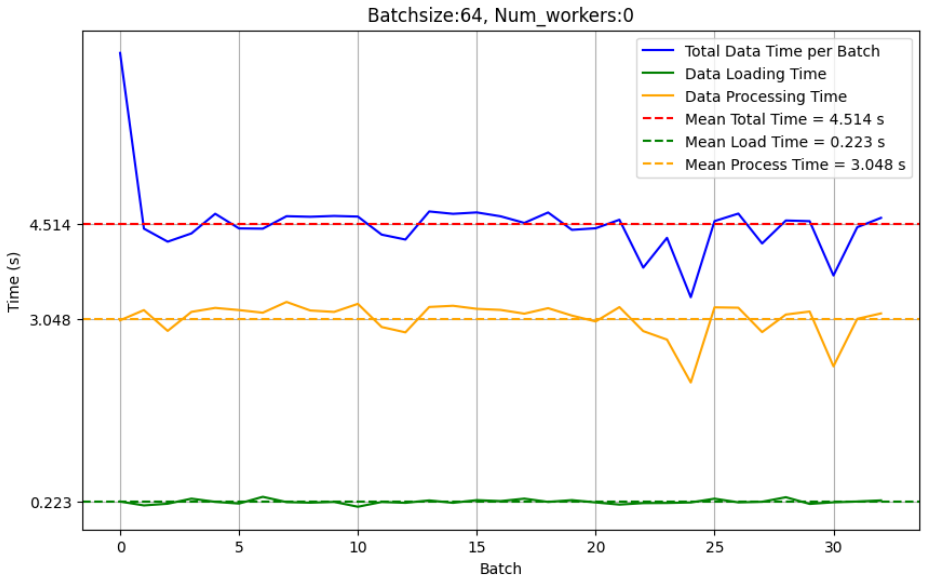
可以看到，8 号机子上的总体偏优，基本比其它两个机子提升两倍之多，我不太了解各服务器的配置，所以暂时不知道具体原因。

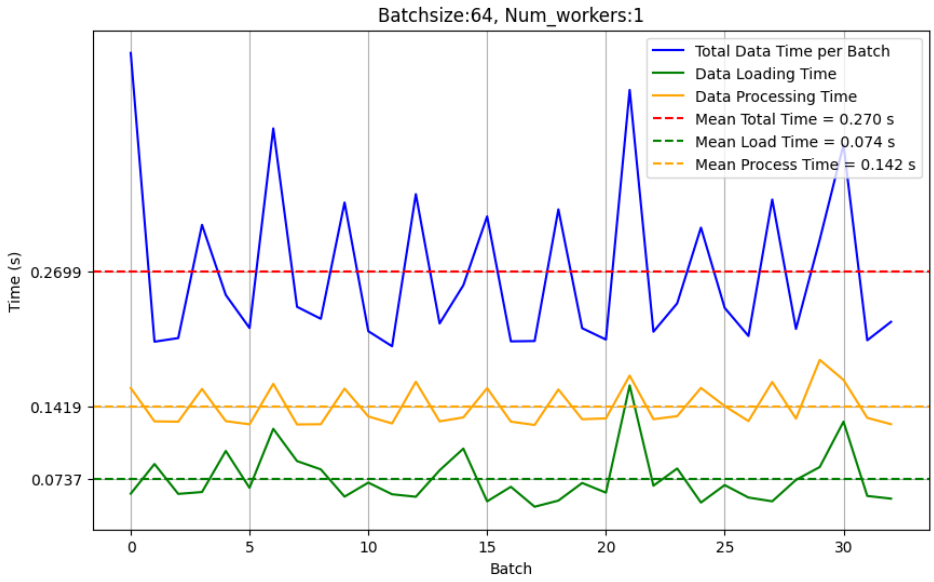
暂时能分析出来：

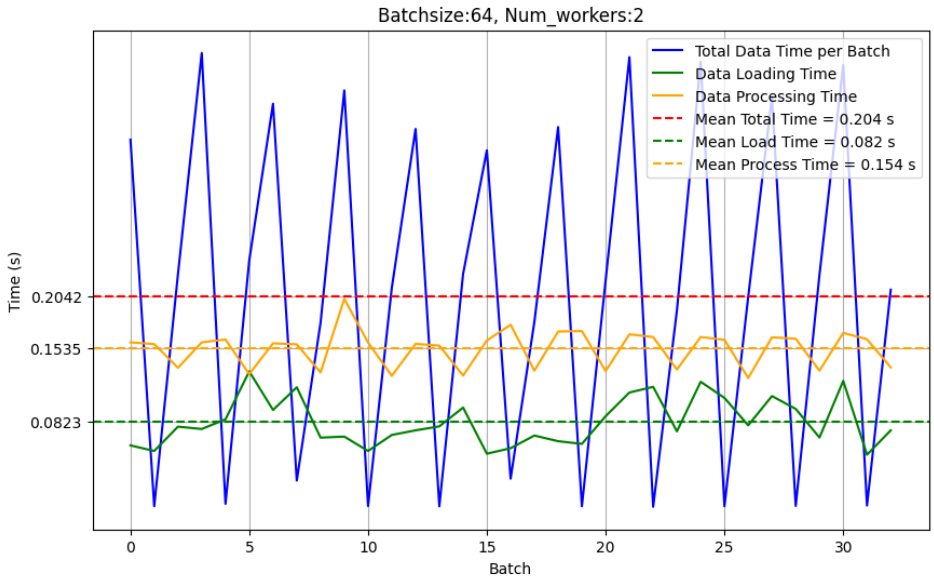
1 同样的数据处理段为什么差别这么大，是因为不同 cpu 的差异吗，毕竟差两倍，在大规模数据集上是 10 小时和 5 小时的区别。

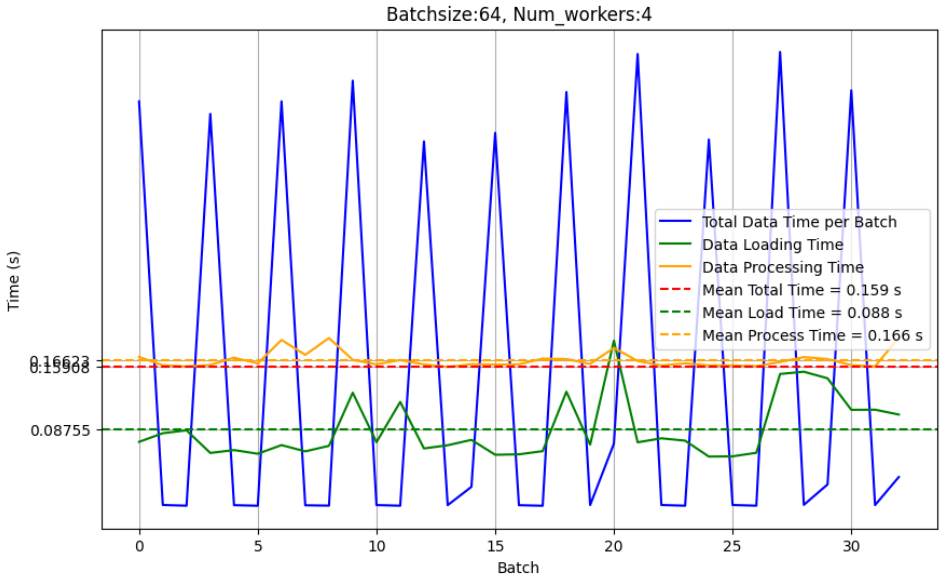
如下是实验截图：

11 号机：

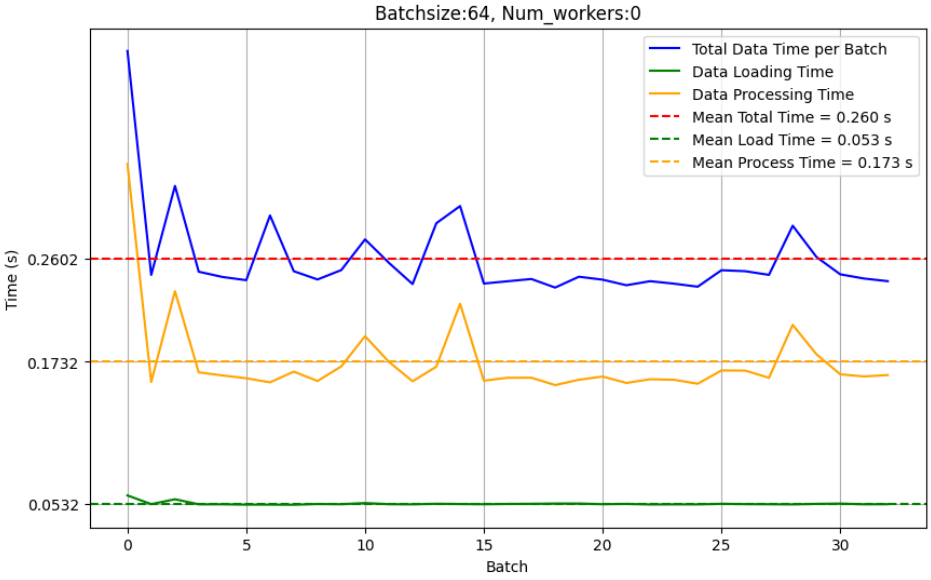


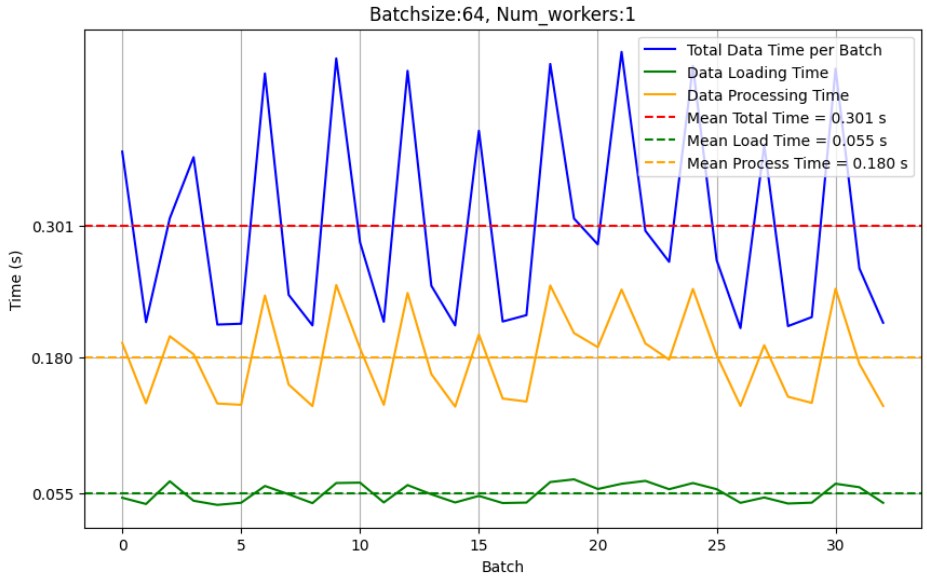

Batchsize:64, Num_workers:1
Total Data Time per Batch
Data Loading Time
Data Processing Time
Mean Total Time = 0.270 s
Mean Load Time = 0.074 s
Mean Process Time = 0.142 s
Time (s)
0.2699
0.1419
0.0737
0
5
10
15
20
25
30
Batch

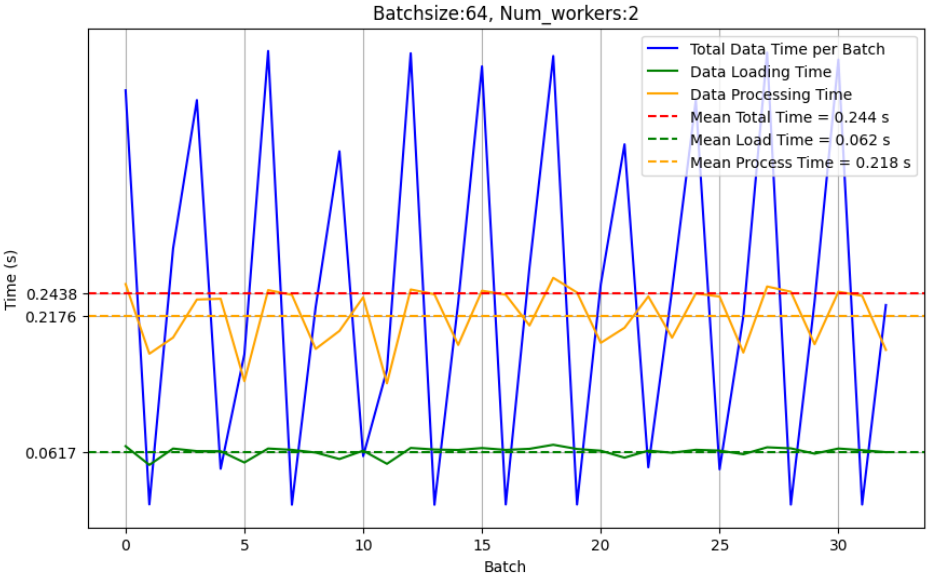

Batchsize:64, Num_workers:2
Total Data Time per Batch
Data Loading Time
Data Processing Time
Mean Total Time = 0.204 s
Mean Load Time = 0.082 s
Mean Process Time = 0.154 s
Time (s)
0.2042
0.1535
0.0823
0
5
10
15
20
25
30
Batch

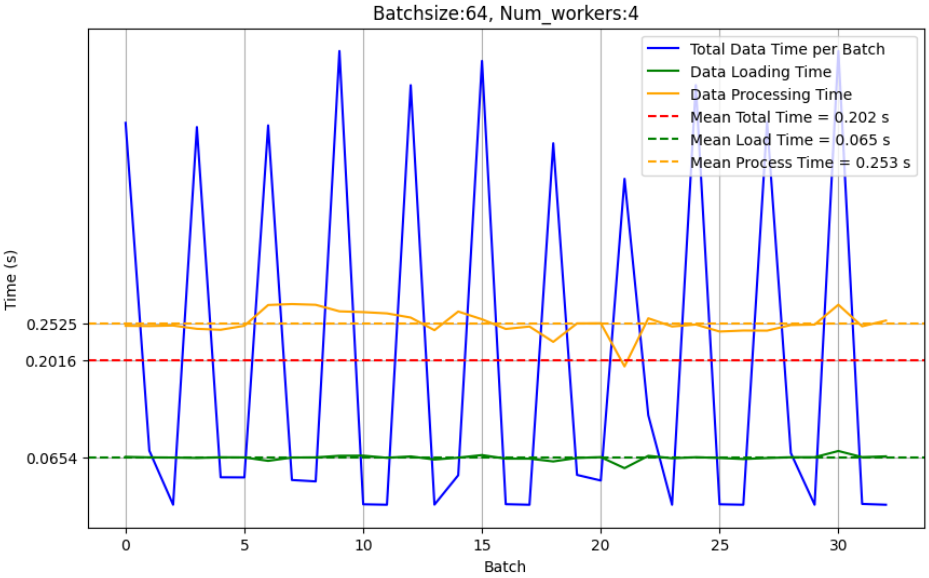

Batchsize:64, Num_workers:4
Total Data Time per Batch
Data Loading Time
Data Processing Time
Mean Total Time = 0.159 s
Mean Load Time = 0.088 s
Mean Process Time = 0.166 s
Time (s)
0.16623
0.15968
0.08755
0
5
10
15
20
25
30
Batch

7 号机：

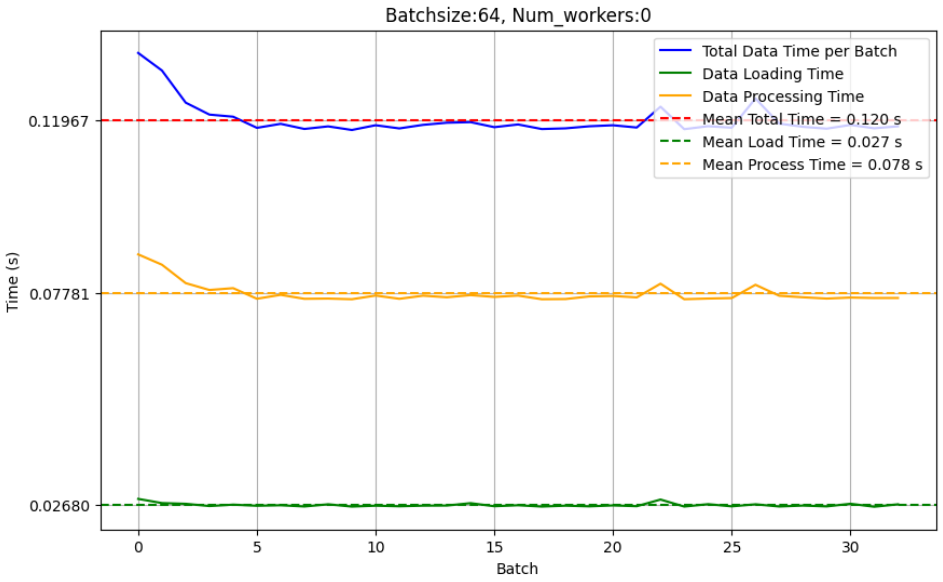

Batchsize:64, Num_workers:0
Total Data Time per Batch
Data Loading Time
Data Processing Time
Mean Total Time = 0.260 s
Mean Load Time = 0.053 s
Mean Process Time = 0.173 s
Time (s)
0.2602
0.1732
0.0532
0
5
10
15
20
25
30
Batch

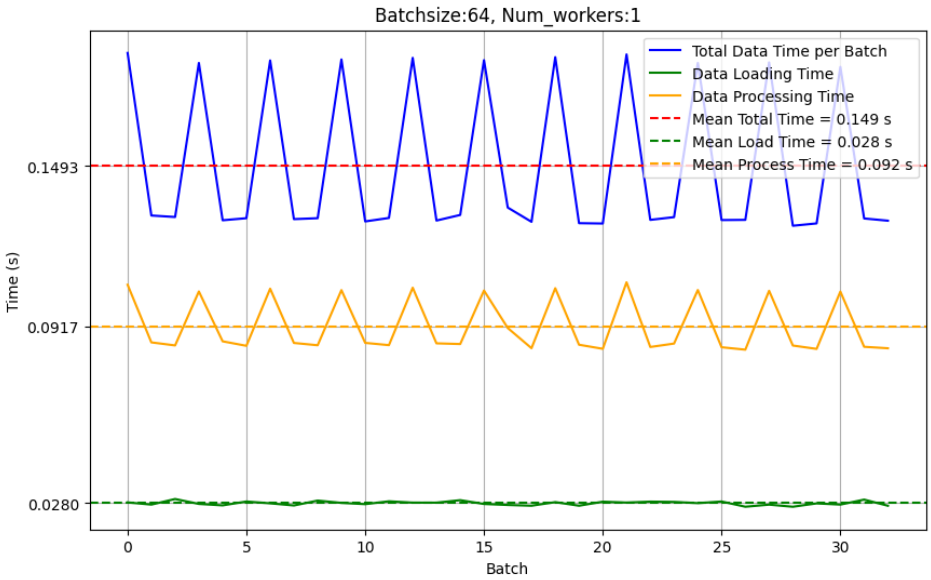

Batchsize:64, Num_workers:1
Total Data Time per Batch
Data Loading Time
Data Processing Time
Mean Total Time = 0.301 s
Mean Load Time = 0.055 s
Mean Process Time = 0.180 s
Time (s)
0.301
0.180
0.055
0
5
10
15
20
25
30
Batch

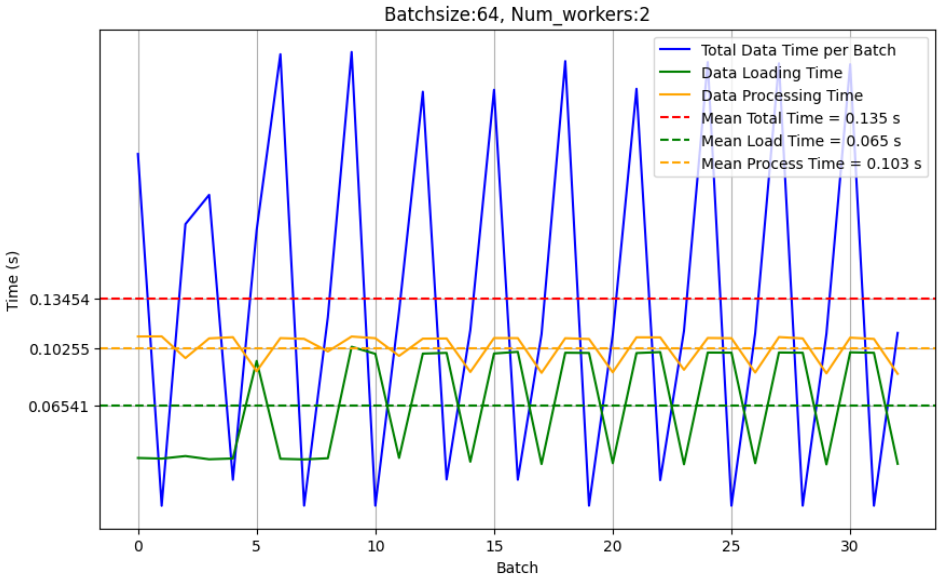

Batchsize:64, Num_workers:2
Total Data Time per Batch
Data Loading Time
Data Processing Time
Mean Total Time = 0.244 s
Mean Load Time = 0.062 s
Mean Process Time = 0.218 s
Time (s)
0.2438
0.2176
0.0617
0
5
10
15
20
25
30
Batch

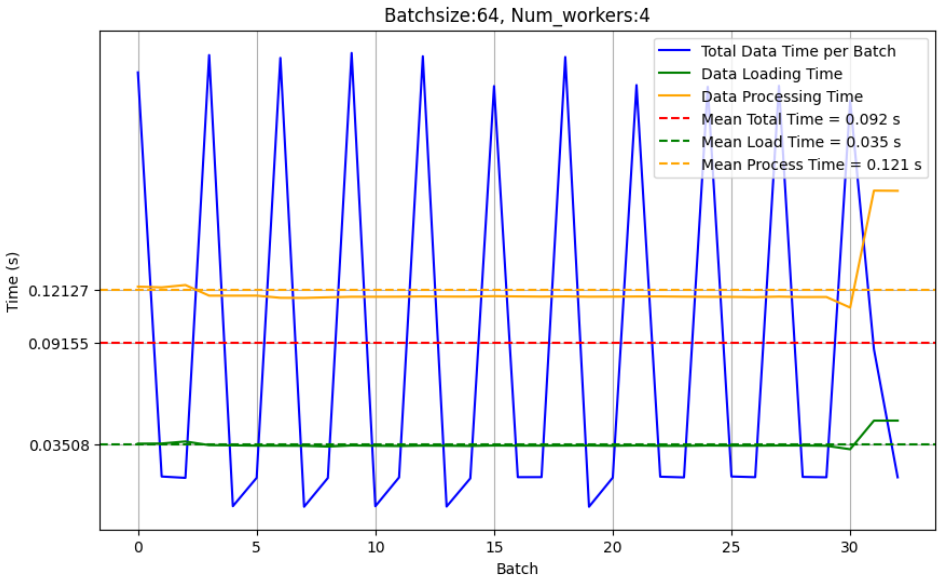

Batchsize:64, Num_workers:4
Total Data Time per Batch
Data Loading Time
Data Processing Time
Mean Total Time = 0.202 s
Mean Load Time = 0.065 s
Mean Process Time = 0.253 s
Time (s)
0.2525
0.2016
0.0654
0
5
10
15
20
25
30
Batch

8 号机：


Batchsize:64, Num_workers:0
Total Data Time per Batch
Data Loading Time
Data Processing Time
Mean Total Time = 0.120 s
Mean Load Time = 0.027 s
Mean Process Time = 0.078 s
0.11967
0.07781
0.02680
Time (s)
0
5
10
15
20
25
30
Batch


Batchsize:64, Num_workers:1
Total Data Time per Batch
Data Loading Time
Data Processing Time
Mean Total Time = 0.149 s
Mean Load Time = 0.028 s
Mean Process Time = 0.092 s
0.1493
0.0917
0.0280
Time (s)
0
5
10
15
20
25
30
Batch


Batchsize:64, Num_workers:2
Total Data Time per Batch
Data Loading Time
Data Processing Time
Mean Total Time = 0.135 s
Mean Load Time = 0.065 s
Mean Process Time = 0.103 s
0.13454
0.10255
0.06541
Time (s)
0
5
10
15
20
25
30
Batch



实际运行：

Num_works=4:

```
- Using Data parallel ...
- [Epoch 0/200][Batch 1/11498] BatchTime = 17.118 (s) DataTime = 12.201 (s) Losses
- [Epoch 0/200][Batch 2/11498] BatchTime = 0.897 (s) DataTime = 0.017 (s) Losses =
- [Epoch 0/200][Batch 3/11498] BatchTime = 0.888 (s) DataTime = 0.006 (s) Losses =
- [Epoch 0/200][Batch 4/11498] BatchTime = 0.888 (s) DataTime = 0.007 (s) Losses =
- [Epoch 0/200][Batch 5/11498] BatchTime = 4.048 (s) DataTime = 3.150 (s) Losses =
- [Epoch 0/200][Batch 6/11498] BatchTime = 1.950 (s) DataTime = 1.065 (s) Losses =
- [Epoch 0/200][Batch 7/11498] BatchTime = 0.887 (s) DataTime = 0.007 (s) Losses =
- [Epoch 0/200][Batch 8/11498] BatchTime = 0.893 (s) DataTime = 0.009 (s) Losses =
- [Epoch 0/200][Batch 9/11498] BatchTime = 7.183 (s) DataTime = 6.283 (s) Losses =
- [Epoch 0/200][Batch 10/11498] BatchTime = 1.202 (s) DataTime = 0.325 (s) Losses =
- [Epoch 0/200][Batch 11/11498] BatchTime = 0.889 (s) DataTime = 0.007 (s) Losses =
- [Epoch 0/200][Batch 12/11498] BatchTime = 0.892 (s) DataTime = 0.017 (s) Losses =
- [Epoch 0/200][Batch 13/11498] BatchTime = 6.075 (s) DataTime = 5.179 (s) Losses =
- [Epoch 0/200][Batch 14/11498] BatchTime = 1.232 (s) DataTime = 0.350 (s) Losses =
- [Epoch 0/200][Batch 15/11498] BatchTime = 0.892 (s) DataTime = 0.008 (s) Losses =
- [Epoch 0/200][Batch 16/11498] BatchTime = 0.897 (s) DataTime = 0.008 (s) Losses =
- [Epoch 0/200][Batch 17/11498] BatchTime = 8.337 (s) DataTime = 7.447 (s) Losses =
- [Epoch 0/200][Batch 18/11498] BatchTime = 1.049 (s) DataTime = 0.167 (s) Losses =
- [Epoch 0/200][Batch 19/11498] BatchTime = 0.891 (s) DataTime = 0.008 (s) Losses =
- [Epoch 0/200][Batch 20/11498] BatchTime = 0.889 (s) DataTime = 0.010 (s) Losses =
- [Epoch 0/200][Batch 21/11498] BatchTime = 7.363 (s) DataTime = 6.464 (s) Losses =
- [Epoch 0/200][Batch 22/11498] BatchTime = 2.696 (s) DataTime = 1.804 (s) Losses =
- [Epoch 0/200][Batch 23/11498] BatchTime = 0.893 (s) DataTime = 0.008 (s) Losses =
- [Epoch 0/200][Batch 24/11498] BatchTime = 0.902 (s) DataTime = 0.015 (s) Losses =
- [Epoch 0/200][Batch 25/11498] BatchTime = 5.857 (s) DataTime = 4.953 (s) Losses =
- [Epoch 0/200][Batch 26/11498] BatchTime = 2.195 (s) DataTime = 1.308 (s) Losses =
```

Num_works:16

```
[Epoch 0/200][Batch 1/11498] BatchTime = 33.885 (s) DataTime = 28.682 (s) Losses = ['0.0000', '137.9634'] lr = 0.000000
[Epoch 0/200][Batch 2/11498] BatchTime = 0.903 (s) DataTime = 0.015 (s) Losses = ['0.0000', '137.8153'] lr = 0.001000
[Epoch 0/200][Batch 3/11498] BatchTime = 0.894 (s) DataTime = 0.009 (s) Losses = ['0.0000', '197.5318'] lr = 0.001000
[Epoch 0/200][Batch 4/11498] BatchTime = 0.899 (s) DataTime = 0.009 (s) Losses = ['0.0000', '149.4078'] lr = 0.001000
[Epoch 0/200][Batch 5/11498] BatchTime = 6.743 (s) DataTime = 5.849 (s) Losses = ['0.0000', '108.9711'] lr = 0.001000
[Epoch 0/200][Batch 6/11498] BatchTime = 1.724 (s) DataTime = 0.844 (s) Losses = ['0.0000', '104.7835'] lr = 0.001000
[Epoch 0/200][Batch 7/11498] BatchTime = 0.903 (s) DataTime = 0.016 (s) Losses = ['0.0000', '88.8002'] lr = 0.001000
[Epoch 0/200][Batch 8/11498] BatchTime = 0.894 (s) DataTime = 0.008 (s) Losses = ['0.0000', '77.5628'] lr = 0.001000
[Epoch 0/200][Batch 9/11498] BatchTime = 0.890 (s) DataTime = 0.008 (s) Losses = ['0.0000', '79.8887'] lr = 0.001000
[Epoch 0/200][Batch 10/11498] BatchTime = 0.891 (s) DataTime = 0.008 (s) Losses = ['0.0000', '80.2600'] lr = 0.001000
[Epoch 0/200][Batch 11/11498] BatchTime = 0.886 (s) DataTime = 0.007 (s) Losses = ['0.0000', '71.8680'] lr = 0.001000
[Epoch 0/200][Batch 12/11498] BatchTime = 0.887 (s) DataTime = 0.007 (s) Losses = ['0.0000', '71.3163'] lr = 0.001000
[Epoch 0/200][Batch 13/11498] BatchTime = 0.886 (s) DataTime = 0.008 (s) Losses = ['0.0000', '70.3161'] lr = 0.001000
[Epoch 0/200][Batch 14/11498] BatchTime = 0.880 (s) DataTime = 0.007 (s) Losses = ['0.0000', '71.2721'] lr = 0.001000
[Epoch 0/200][Batch 15/11498] BatchTime = 0.885 (s) DataTime = 0.007 (s) Losses = ['0.0000', '70.8956'] lr = 0.001000
[Epoch 0/200][Batch 16/11498] BatchTime = 0.904 (s) DataTime = 0.019 (s) Losses = ['0.0000', '70.8994'] lr = 0.001000
[Epoch 0/200][Batch 17/11498] BatchTime = 10.369 (s) DataTime = 9.479 (s) Losses = ['0.0000', '63.0775'] lr = 0.001000
[Epoch 0/200][Batch 18/11498] BatchTime = 5.434 (s) DataTime = 4.542 (s) Losses = ['0.0000', '63.0981'] lr = 0.001000
[Epoch 0/200][Batch 19/11498] BatchTime = 0.890 (s) DataTime = 0.008 (s) Losses = ['0.0000', '62.1150'] lr = 0.001000
[Epoch 0/200][Batch 20/11498] BatchTime = 0.891 (s) DataTime = 0.008 (s) Losses = ['0.0000', '61.7713'] lr = 0.001000
[Epoch 0/200][Batch 21/11498] BatchTime = 10.413 (s) DataTime = 9.524 (s) Losses = ['0.0000', '65.3981'] lr = 0.001000
[Epoch 0/200][Batch 22/11498] BatchTime = 3.658 (s) DataTime = 2.752 (s) Losses = ['0.0000', '66.5505'] lr = 0.001000
[Epoch 0/200][Batch 23/11498] BatchTime = 0.895 (s) DataTime = 0.010 (s) Losses = ['0.0000', '67.6510'] lr = 0.001000
[Epoch 0/200][Batch 24/11498] BatchTime = 0.887 (s) DataTime = 0.009 (s) Losses = ['0.0000', '65.0815'] lr = 0.001000
[Epoch 0/200][Batch 25/11498] BatchTime = 0.889 (s) DataTime = 0.007 (s) Losses = ['0.0000', '63.0594'] lr = 0.001000
[Epoch 0/200][Batch 26/11498] BatchTime = 0.884 (s) DataTime = 0.008 (s) Losses = ['0.0000', '64.2656'] lr = 0.001000
[Epoch 0/200][Batch 27/11498] BatchTime = 2.786 (s) DataTime = 1.892 (s) Losses = ['0.0000', '63.4628'] lr = 0.001000
[Epoch 0/200][Batch 28/11498] BatchTime = 0.895 (s) DataTime = 0.013 (s) Losses = ['0.0000', '62.3571'] lr = 0.001000
[Epoch 0/200][Batch 29/11498] BatchTime = 0.892 (s) DataTime = 0.008 (s) Losses = ['0.0000', '63.9437'] lr = 0.001000
[Epoch 0/200][Batch 30/11498] BatchTime = 0.887 (s) DataTime = 0.008 (s) Losses = ['0.0000', '61.0530'] lr = 0.001000
[Epoch 0/200][Batch 31/11498] BatchTime = 0.894 (s) DataTime = 0.008 (s) Losses = ['0.0000', '60.0711'] lr = 0.001000
[Epoch 0/200][Batch 32/11498] BatchTime = 0.901 (s) DataTime = 0.014 (s) Losses = ['0.0000', '58.8759'] lr = 0.001000
[Epoch 0/200][Batch 33/11498] BatchTime = 11.617 (s) DataTime = 10.709 (s) Losses = ['0.0000', '63.6782'] lr = 0.001000
[Epoch 0/200][Batch 34/11498] BatchTime = 4.482 (s) DataTime = 3.581 (s) Losses = ['0.0000', '58.2843'] lr = 0.001000
[Epoch 0/200][Batch 35/11498] BatchTime = 0.894 (s) DataTime = 0.009 (s) Losses = ['0.0000', '60.3085'] lr = 0.001000
[Epoch 0/200][Batch 36/11498] BatchTime = 0.889 (s) DataTime = 0.009 (s) Losses = ['0.0000', '64.3093'] lr = 0.001000
[Epoch 0/200][Batch 37/11498] BatchTime = 10.303 (s) DataTime = 9.376 (s) Losses = ['0.0000', '65.6142'] lr = 0.001000
[Epoch 0/200][Batch 38/11498] BatchTime = 2.136 (s) DataTime = 1.236 (s) Losses = ['0.0000', '61.2851'] lr = 0.001000
[Epoch 0/200][Batch 39/11498] BatchTime = 0.918 (s) DataTime = 0.011 (s) Losses = ['0.0000', '53.7125'] lr = 0.001000
[Epoch 0/200][Batch 40/11498] BatchTime = 1.578 (s) DataTime = 0.677 (s) Losses = ['0.0000', '59.6389'] lr = 0.001000
[Epoch 0/200][Batch 41/11498] BatchTime = 0.903 (s) DataTime = 0.011 (s) Losses = ['0.0000', '60.8825'] lr = 0.001000
[Epoch 0/200][Batch 42/11498] BatchTime = 0.884 (s) DataTime = 0.009 (s) Losses = ['0.0000', '66.3812'] lr = 0.001000
[Epoch 0/200][Batch 43/11498] BatchTime = 0.893 (s) DataTime = 0.008 (s) Losses = ['0.0000', '58.0004'] lr = 0.001000
```