C1.

```
Epoch 0:
 Loss Mean: 0.4652
 Precision@1: 0.775
 Precision@3: 0.516
Epoch 1:
 Loss Mean: 0.2692
 Precision@1: 0.926
 Precision@3: 0.639
Epoch 2:
 Loss Mean: 0.2501
 Precision@1: 0.926
 Precision@3: 0.643
Epoch 3:
 Loss Mean: 0.2447
 Precision@1: 0.926
 Precision@3: 0.644
Epoch 4:
 Loss Mean: 0.2391
 Precision@1: 0.926
 Precision@3: 0.645
```

C2:

```
Average Precision@1:0.8958
Average Precision@3:0.6174
Total Time for Epochs: 1381.817372737 secs
Total Time for data loading: 1170.4243495264 secs
Total Time for batch execution: 1375.2644689786 secs
Average Time for an epoch: 276.36347454739735 secs
Average Time for data loading: 1.9507072492 secs
Average Time for one batch execution: 2.292107448297708 secs
```

C3:

Workers	Avg Data Loading Time secs	Total Data Loading Time secs
0	2.349345622	1409.607373
1	1.920952091	1152.571254
2	0.9981902984	598.914179
4	0.3486322972	209.1793783
8	0.02762076367	16.5724582
12	0.02986814922	17.92088953
16	0.0308063115	18.4837869
20	0.05072663657	30.43598194

24	0.07466173445	44.79704067
28	0.06794440018	40.76664011

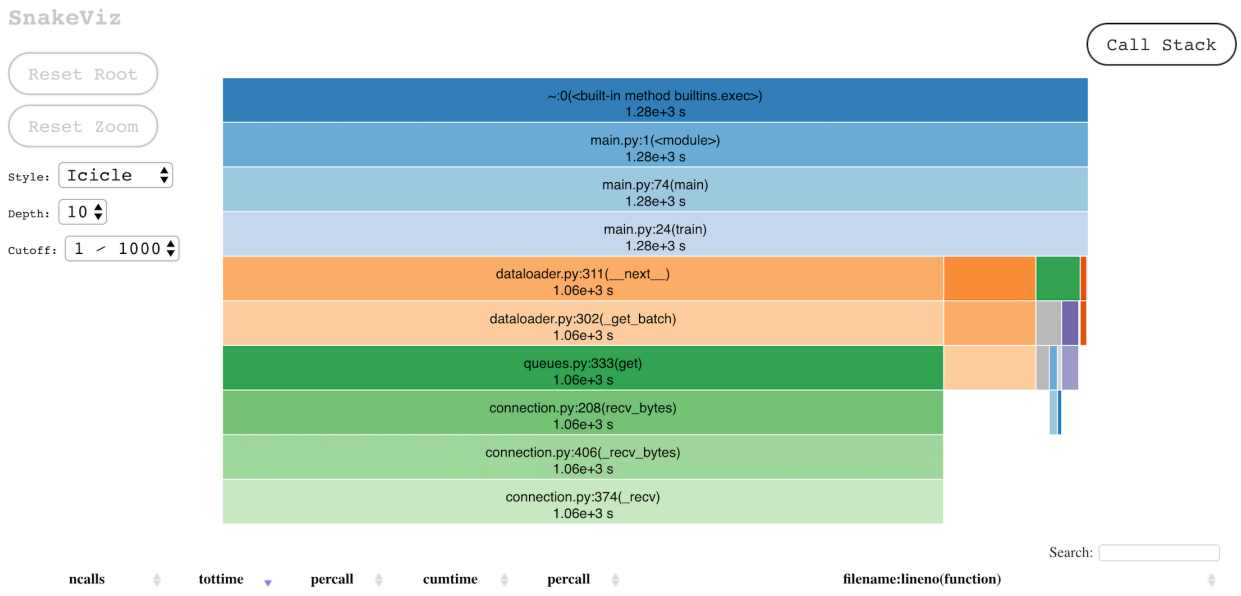
Best number of workers : 8

C4:
Num_Workers : 1

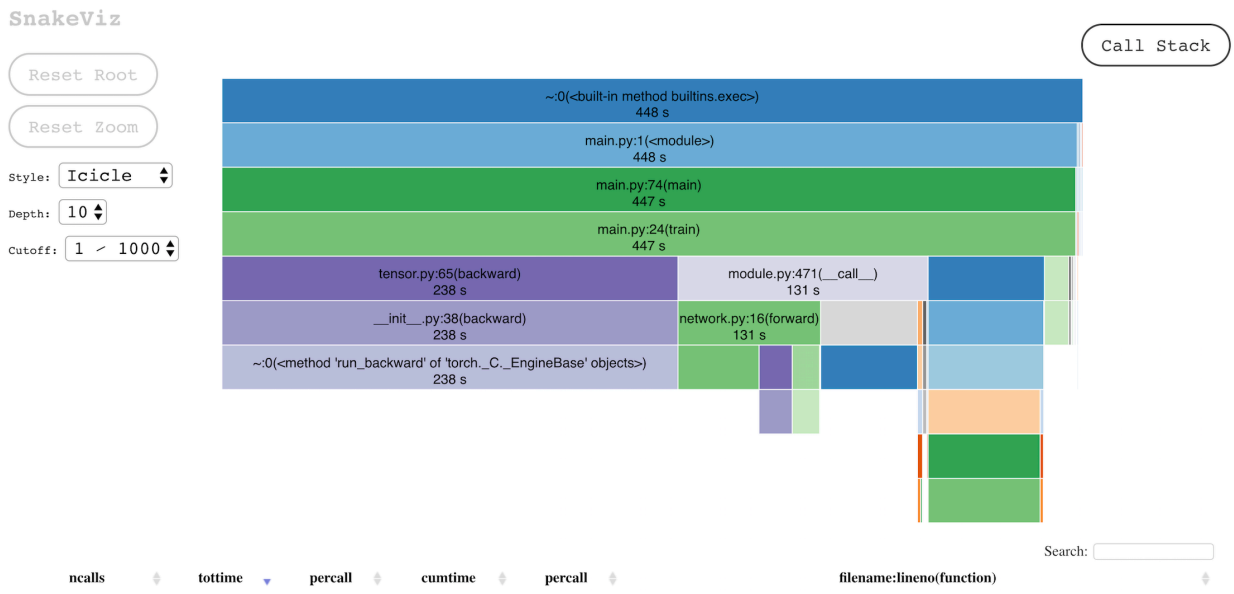
```
Average Time for an epoch: 255.58068878008052 secs
Average Time for data loading: 1.7726490107 secs
Total Time for data loading: 1063.5894064009 secs
Average Time for one batch execution: 2.1134020699506313 secs
```

```
Num_Workers : 8
Average Time for an epoch: 89.35773856686428 secs
Average Time for data loading: 0.1034247921 secs
Total Time for data loading: 62.0548752663 secs
Average Time for one batch execution: 0.7233552696734357 secs
```

Plots:
Num_Workers : 1



Num_Workers : 8



Explanation:
As the number of workers increases, more parallel threads can work simultaneously to load the data so the time reduces. But the overhead of creating threads takes over the reduced time as the number of threads keeps increasing.

C5:

CPU with 1 worker:
Average Precision@1:0.8958
Average Precision@3:0.6174
Average Time for an epoch: 276.36347454739735 secs
Total Time for data loading: 1170.4243495264 secs
Average Time for one batch execution: 2.292107448297708 secs

GPU with 1 worker:
Average Precision@1:0.8916
Average Precision@3:0.6228
Average Time for an epoch: 410.29961487525145 secs
Total Time for data loading: 1936.6304472878 secs

Average Time for one batch execution: 3.2371580140312046 secs

SGD with 8 workers:
Average loss: 0.27972
Average Precision@1:0.9038
Average Precision@3:0.6488
Average Time for an epoch: 263.7599270750303 secs
Total Time for data loading: 1165.4430829538 secs
Average Time for one batch execution: 1.96510496428508 secs

Nesterov with 8 workers:
Average loss: 0.28238
Average Precision@1:0.9114
Average Precision@3:0.6322
Average Time for an epoch: 256.04222603761593 secs
Total Time for data loading: 1132.3305943795 secs
Average Time for one batch execution: 1.909445991657752 secs

Adam with 8 workers:
Average loss: 0.21428
Average Precision@1:0.9398
Average Precision@3:0.664
Average Time for an epoch: 247.08243355820886 secs
Total Time for data loading: 1061.6790117347 secs
Average Time for one batch execution: 1.796133587200893 secs

Adadelta with 8 workers:
Average loss: 0.23202
Average Precision@1:0.9294
Average Precision@3:0.6482
Average Time for an epoch: 106.76676436811685 secs
Total Time for data loading: 357.8039726326 secs
Average Time for one batch execution: 0.6249504870533322 secs

Adagrad with 8 workers:
Average loss: 0.21498
Average Precision@1:0.9374
Average Precision@3:0.6624
Average Time for an epoch: 206.38922340483404 secs
Total Time for data loading: 780.516976265 secs
Average Time for one batch execution: 1.3386395732755774 secs