Assignments around CNN scaling

During the course we have used a Horovod-enabled CIFAR10 distributed training implementation. We have performed ResNet-56 training using 8 Broadwell nodes from Cartesius and 2 processes on each of the nodes, leading to 16 workers in total. Each process was working with a batch size of 128 images and hence the total (global) batch size was 2048. We have performed 80 epochs of training, meaning 1953 SGD iterations (80 epochs \* 50,000 images / 2048). To compensate for the large batch size (2048), we have linearly increased the learning rate to account for the larger global batch size.

1. As assignment, we would like you to extend the following MNIST codebase with Horovod support: <https://github.com/tensorflow/models/tree/master/official/mnist>. The code uses a default batch size of 100 and 40 epochs (tip: read the ‘Usage’ section on https://github.com/uber/horovod).
2. How many iterations should be performed to cover the 40 epochs with a batch size of 100 (the number of training images is 60,000)? What is the time required for a full training run on a Broadwell CPU node?
3. Run the resulting code MNIST on a single node, but with various settings for the number of processes per node. What number of processes per node behaves best in terms of throughput? Also, do not forget to alter the inter\_op/intra\_op parameters. Which options give highest throughput?
4. Run the resulting Horovod code on 2/4/8 nodes with the best number of processes per node and inter\_op/intra\_op settings that you’ve discovered in the previous exercise. Could you plot the throughput for each of the runs? What is the scaling efficiency when going from 1 node to 8 nodes? What could be the potential bottlenecks?
5. Profile your communication with HOROVOD\_TIMELINE. What is the bottleneck? More details on how to use HOROVOD\_TIMELINE can be found here: https://github.com/uber/horovod/blob/master/docs/timeline.md
6. What is the validation accuracy for the 1/2/4/8 node versions?
7. What happens to the batch size and learning rate? How can we keep the validation accuracy high when scaling up the batch size?
8. What is the best validation accuracy that you have achieved using 8 nodes, and the same 40 training epochs?

Your accounts will be valid for a week, and we advise you to use the Broadwell CPU partition for all experiments.