How long does it take to complete the training run?

Over 2 days for max_steps = 100,000 (on v100s)

Do you think your model is fully trained? How can you tell?

Yes. Training loss has leveled out.

Were you overfitting?

No.

Were your GPUs fully utilized?

Yes.

Did you monitor network traffic (hint: apt install nmon)? Was network the bottleneck?

I did not.

Take a look at the plot of the learning rate and then check the config file. Can you explain this setting?

Learning rate is 2.0 with 8000 steps. As the model gets closer to the target predictions, we need to reduce the learning rate.

How big was your training set (mb)? How many training lines did it contain?

train.clean.en.shuffled = ~950MB
train.clean.de.shuffled = ~1000MB
~4.5M lines

What are the files that a TF checkpoint is comprised of?

- Checkpoint file
- Data
- Index file
- Meta file

How big is your resulting model checkpoint (mb)?

Under 1MB

Remember the definition of a "step". How long did an average step take?

Average step is ~ 1.5 seconds

How does that correlate with the observed network utilization between nodes?

This is max speed as the GPUs were fully utilized.