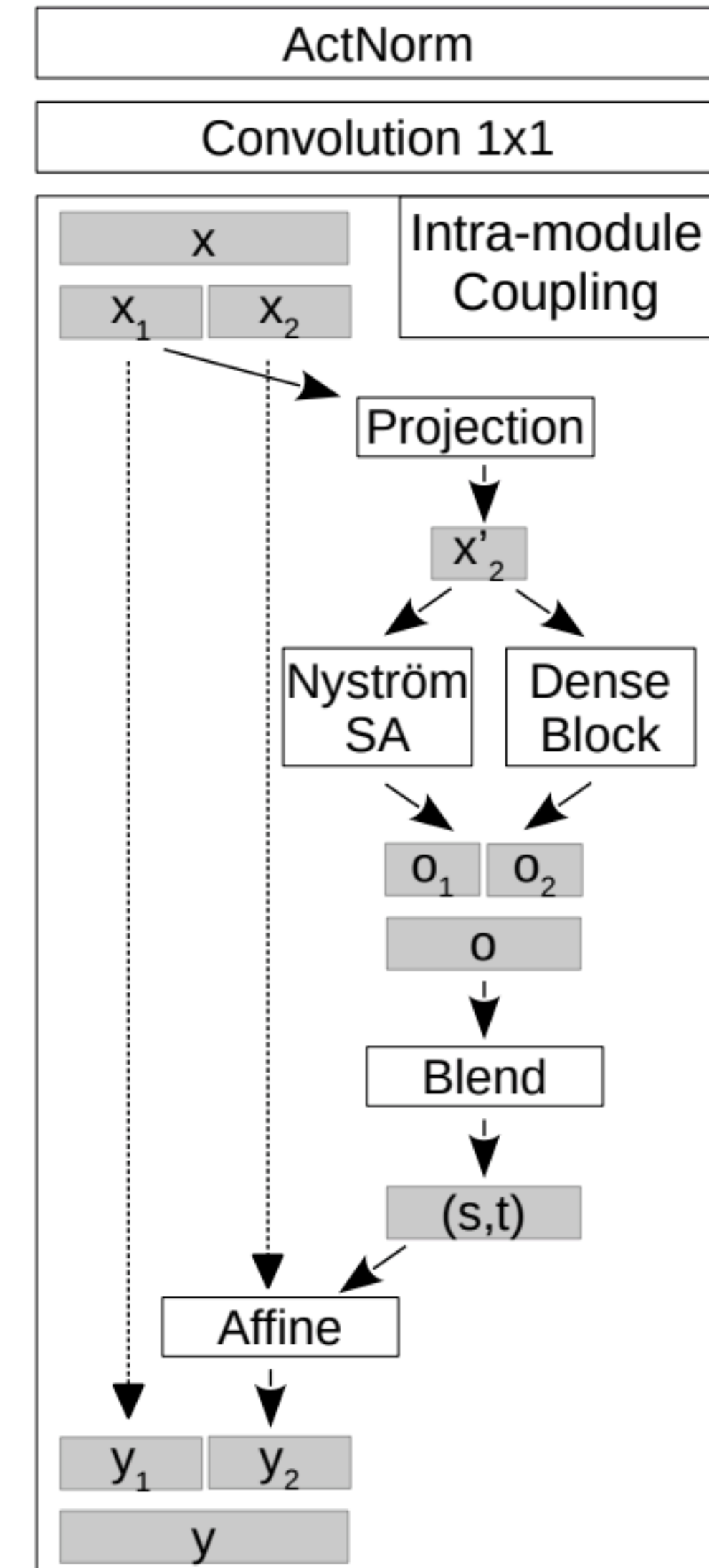


Efficient training

GLOW results took 40 GPU weeks to train.

Transformer based Normalizing Flow gets similar performance with 1 GPU week.



Normalizing Flows

Extensions: Conditional Sampling