Training loss: workers = 4,  $\tau$  = 8,  $\beta$  = 3.96 2.0 Mean loss 1.0 0.5 0.0 200 400 600 Wall-clock time (s)