a9a; heterogeneous; batch = 20; workers = 100 Local SGD 10 local steps ★ Local SGD 20 local steps Local SGD 30 local steps 10^{0} Local SGD 50 local steps Local SGD 100 local steps $||\nabla f(x)||$ 10^{-1} 25 50 75 125 150 100 175 200 0 communications