Let us recall parallel gradient descent

Server

Param

The i-th worker performs:

1. Receiving model parameters **w** from the server.

2. Using **w** and its local data to compute gradient $\mathbf{g}_i$.

3. Sending $\mathbf{g}_i$ to the server.