

- ightharpoonup Gradient  $abla_t$  on Node k
- Accumulated Gradient  $v_t$  on Node k
- ightharpoonup Momentum  $u_{t-1}$  on Server
- Update Velocity  $\Delta = u_t$  on Server
- **B** Original Optimization Direction
- C Optimization DirectionWithout Momentum Correction