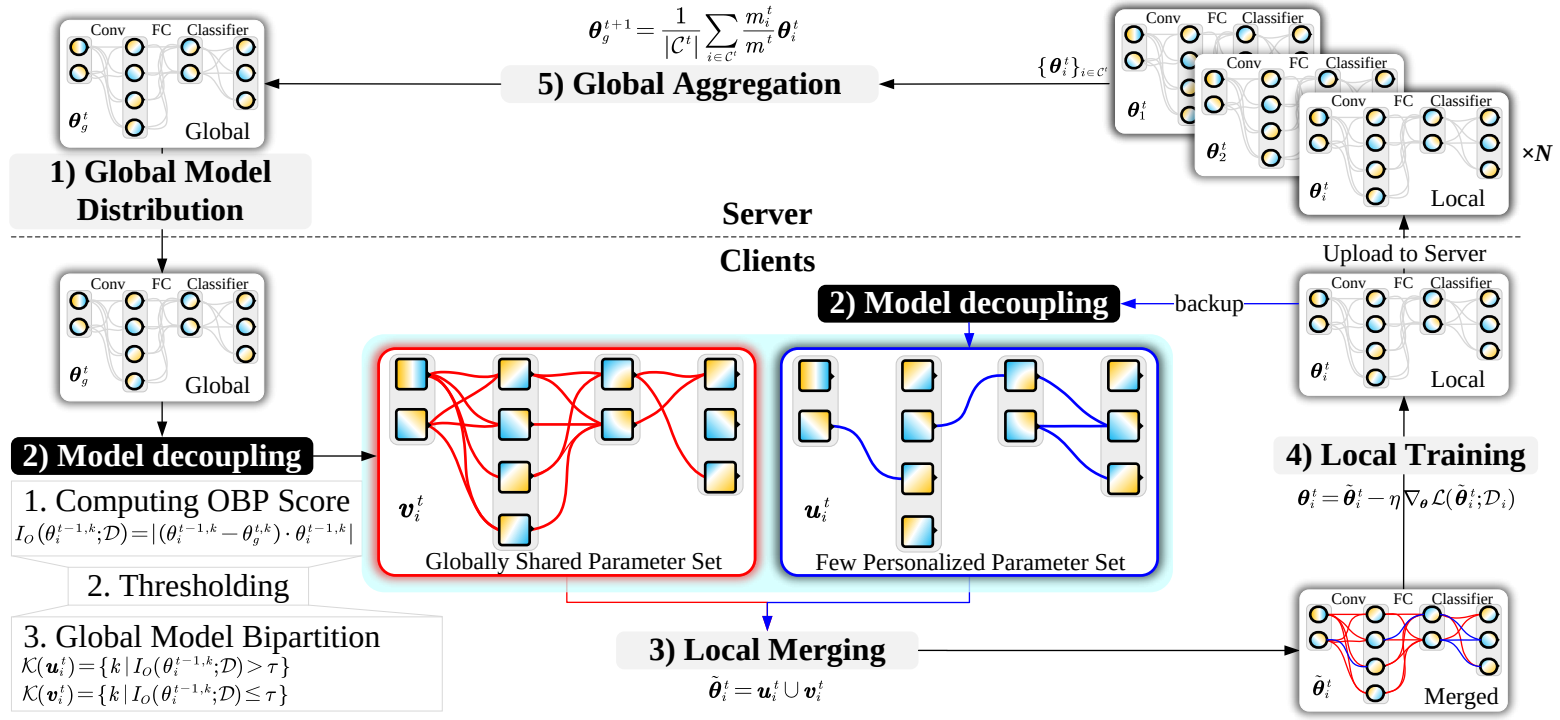
$\boldsymbol{\theta}_g^{t+1} = \frac{1}{|\mathcal{C}^t|} \sum_{i \in \mathcal{C}^t} \frac{m_i^t}{m^t} \boldsymbol{\theta}_i^t$
5) Global Aggregation
$\{\boldsymbol{\theta}_i^t\}_{i \in \mathcal{C}^t}$
Conv FC Classifier
Global
Local
$\times N$
1) Global Model Distribution
Server
Clients
Upload to Server
backup
2) Model decoupling
2) Model decoupling
1. Computing OBP Score
$I_O(\theta_i^{t-1,k}; \mathcal{D}) = |(\theta_i^{t-1,k} - \theta_g^{t,k}) \cdot \theta_i^{t-1,k}|$
2. Thresholding
3. Global Model Bipartition
$\mathcal{K}(\boldsymbol{u}_i^t) = \{k \mid I_O(\theta_i^{t-1,k}; \mathcal{D}) > \tau\}$
$\mathcal{K}(\boldsymbol{v}_i^t) = \{k \mid I_O(\theta_i^{t-1,k}; \mathcal{D}) \leq \tau\}$
Globally Shared Parameter Set
Few Personalized Parameter Set
4) Local Training
$\boldsymbol{\theta}_i^t = \tilde{\boldsymbol{\theta}}_i^t - \eta \nabla_{\boldsymbol{\theta}} \mathcal{L}(\tilde{\boldsymbol{\theta}}_i^t; \mathcal{D}_i)$
3) Local Merging
$\tilde{\boldsymbol{\theta}}_i^t = \boldsymbol{u}_i^t \cup \boldsymbol{v}_i^t$
Merged