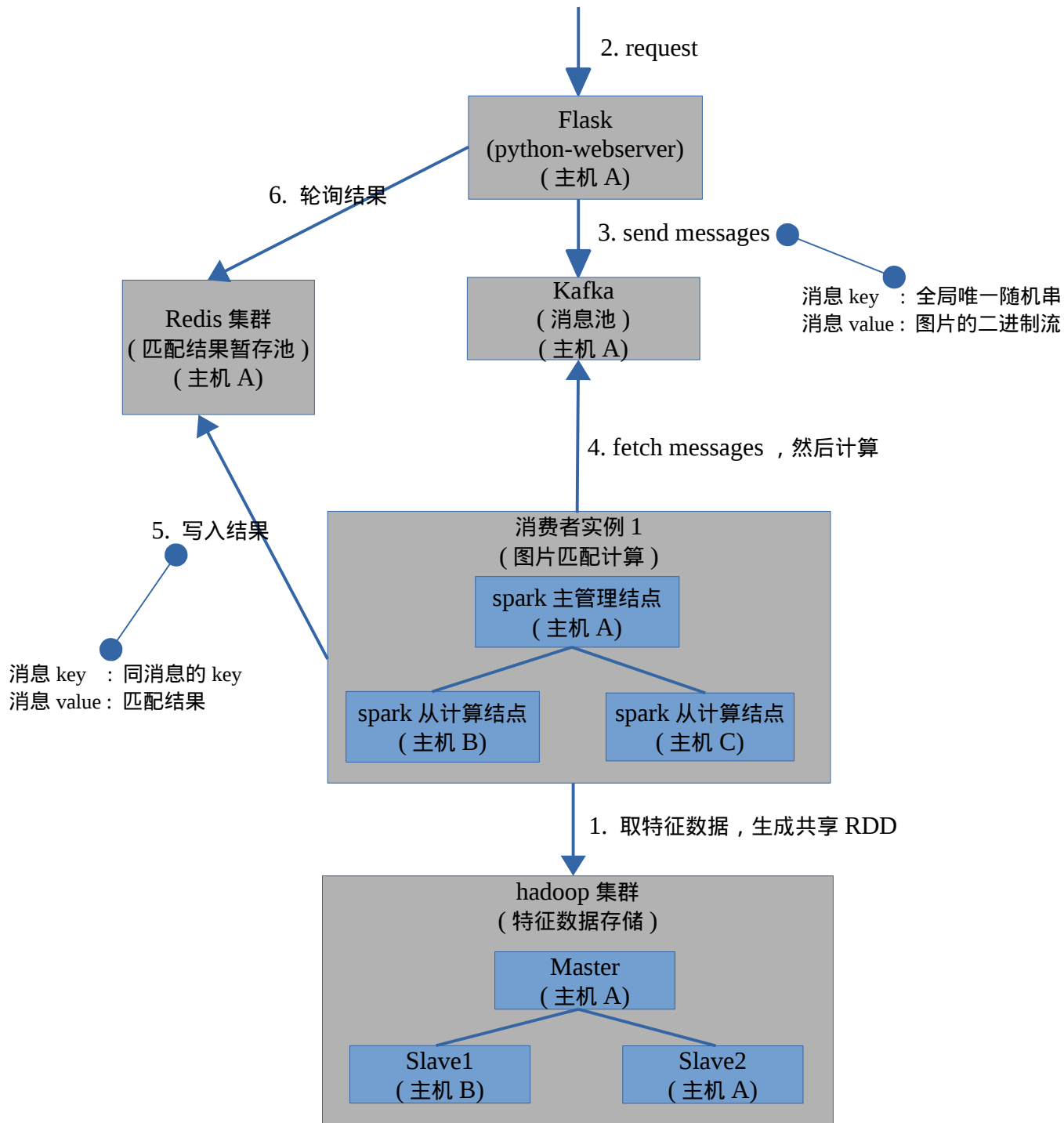


注：消费者不能是 spark-streaming，原因在 mapreduce.md 中有说明
注：关于当前架构的部署，见于 deploy.md

图片识别业务的架构方案

一. 当前架构（3 台主机）



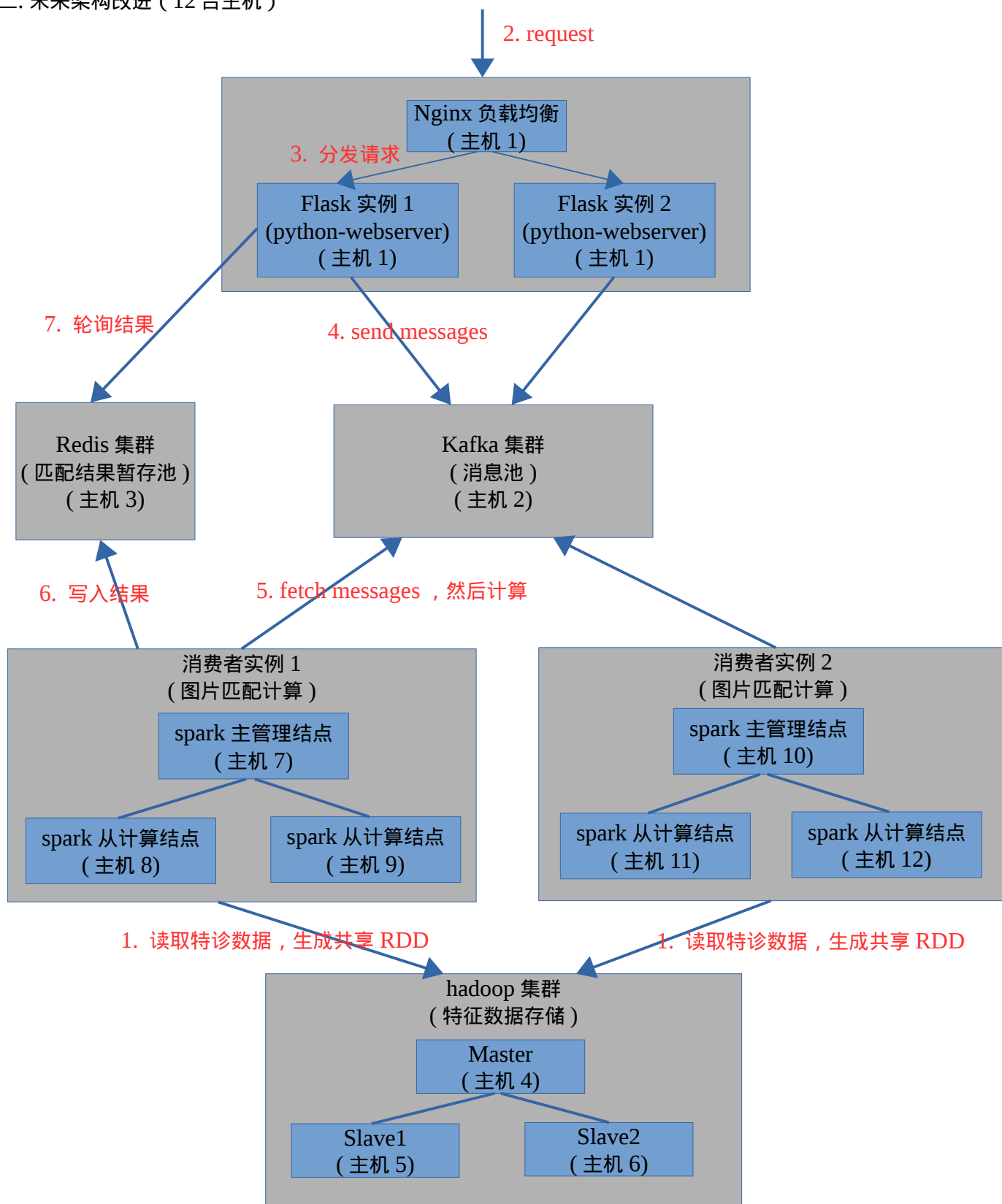
时序

启动 spark：1

处理请求：2 → 3 → 4 → 5 (主分支)

→ 6 (另一条分支，即从 3 开始，两条分支并行)

二. 未来架构改进 (12 台主机)



时序

启动 spark : 1

处理请求 : 2 → 3 → 4 → 5 → 6 (主分支)
→ 7 (另一条分支, 即从 3 开始, 两条分支并行)