

分布式云平台

讲师：肖斌



- 问题：
 - 在QQ、facebook、人人网等社交网站上，为了使用户在网络上认识更多的朋友，社交网站往往提供类似“你可能感兴趣的人”、“你可能认识的人”、“间接关注推荐”等好友推荐的功能。一直很好奇这个功能是怎么实现的。
- 环境
 - Hadoop2.5_x64版本
 - Eclipse3.5
 - Jdk1.7



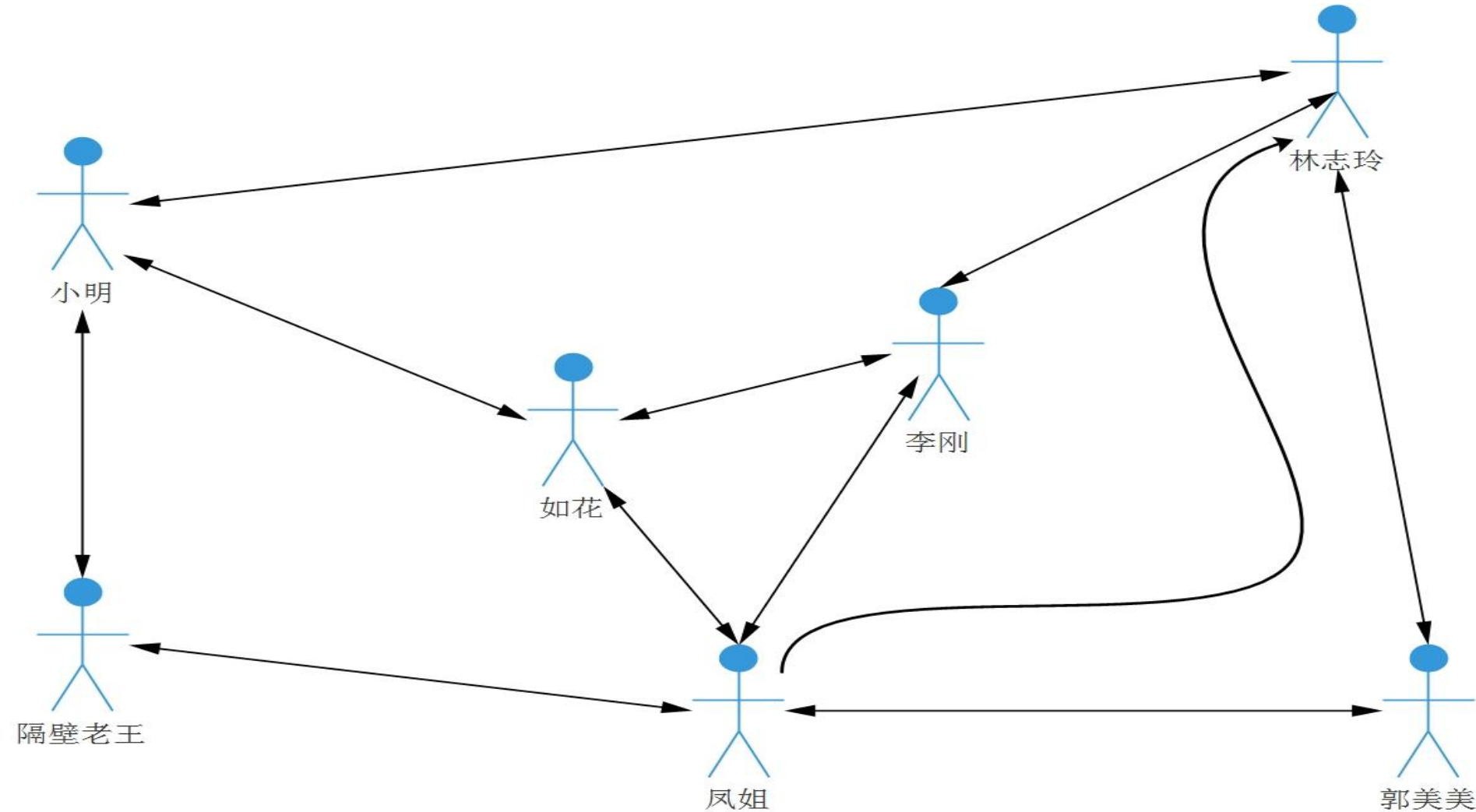
- Hadoop2.5安装（伪分布式）步骤
 - 1、检查环境
 - JDK是否安装，环境变量是否正常
 - 2、本地免密码登录
 - `ssh-keygen -t dsa -P "" -f ~/.ssh/id_dsa`
 - `cat ~/.ssh/id_dsa.pub >> ~/.ssh/authorized_keys`
 - 3、解压hadoop
 - `tar -zxvf hadoop-2.x.x_x64.tar.gz`
 - 4、配置hadoop的配置文件
 - `hadoop-env.sh`
 - `hdfs-site.xml`
 - `mapred-site.xml`
 - `yarn-site.xml`
- 具体可以参考
 - <http://hadoop.apache.org/docs/r2.5.2/hadoop-project-dist/hadoop-common/SingleCluster.html>



- 配置hadoop的环境变量
 - vi ~/.bash_profile
 - source ~/.bash_profile
- 启动hadoop
 - start-all.sh
- 检查hadoop是否已经成功启动
 - 1、检查hdfs : <http://node9:50070>
 - 2、检查Mapreduce : <http://node9:8088>



- 社交网站上的各个用户以及用户之间的相互关系图



- Mapreduce思路

- 找出每一对直接（一度关系）好友
- 找出每一对好友的（二度关系）好友
- 从（二度关系）好友对中去掉那些已经是直接好友的
- 并且要统计二度关系好友的支持数
 - 支持数：二度关系好友之间可以通过多少个相互直接好友认识的。
- 按照支持数排序



- 执行Mapreduce的环境



- 课程大纲
- **三大计算框架**
- 服务器分布式，集群，虚拟化技术。

