



大数据开源平台与工具

北京理工大学计算机学院 王一拙

2019年1月



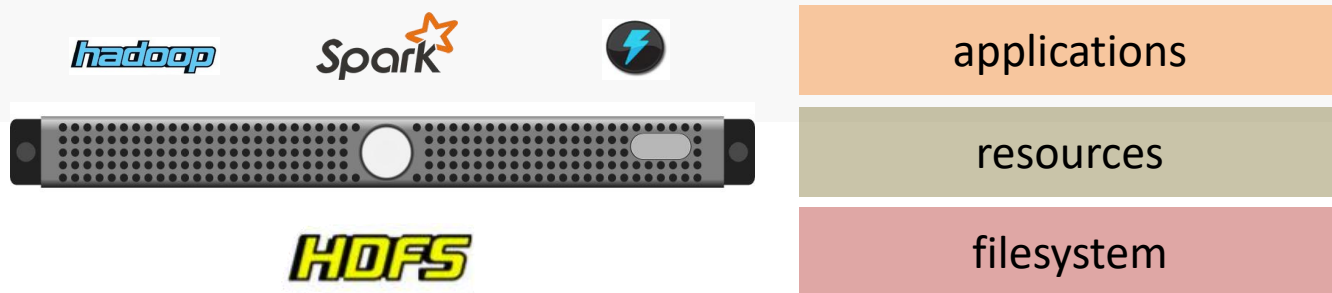
内容提要

- 数据采集与清洗
- 数据存储与管理
- 数据处理与分析
- 资源管理与调度



4. 资源管理与调度

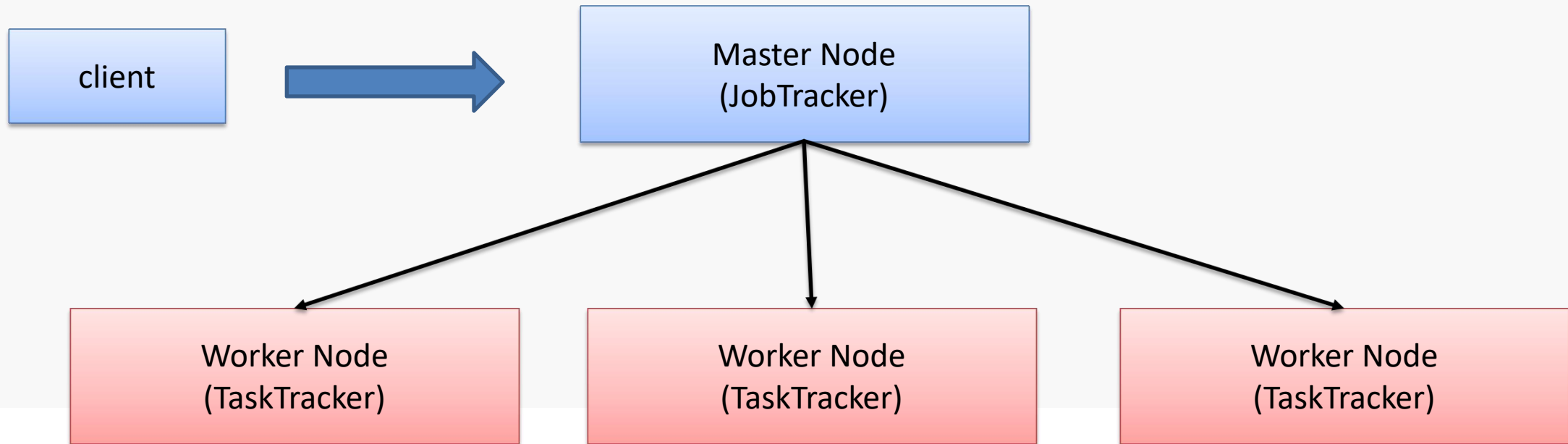
- 什么是资源管理与调度
- 资源管理软件
 - 开源软件 Apache YARN和Mesos
 - 商业软件 Google Borg和Omega
- YARN资源管理框架
- Zookeeper集群节点协调





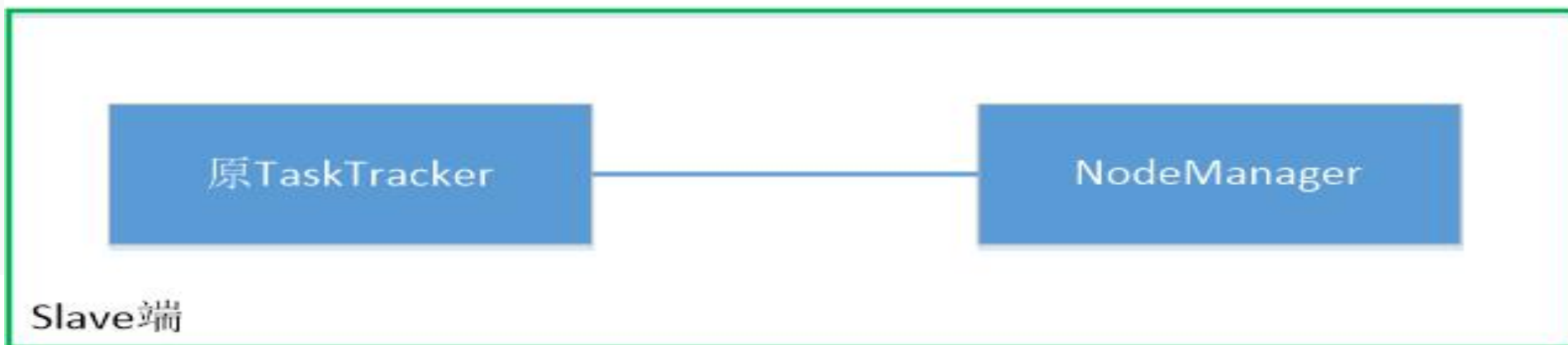
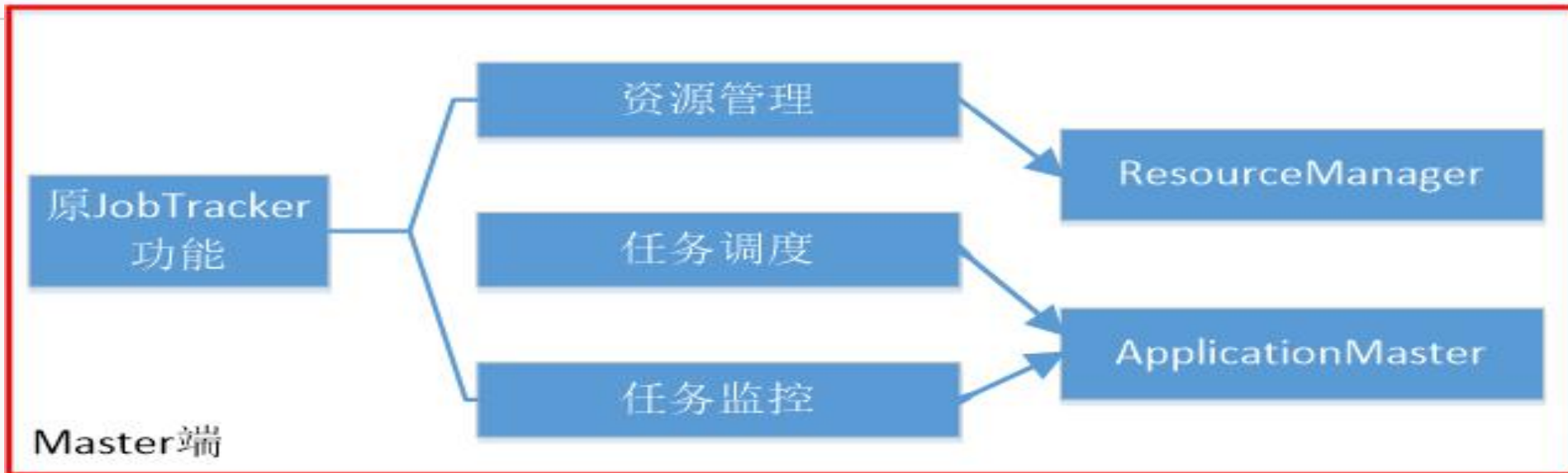
Hadoop 1.0

- Master: JobTracker (JT)
- Worker: Tasktracker (TT)
 - Fixed # of map slots and reduce slots





YARN设计思路





- 处理客户端请求
- 启动/监控ApplicationMaster
- 监控NodeManager
- 资源分配与调度

- 为应用程序申请资源，并分配给内部任务
- 任务调度、监控与容错

- 单个节点上的资源管理
- 处理来自ResourceManger的命令
- 处理来自ApplicationMaster的命令





Zookeeper



Zookeeper概述

- » Zookeeper 是 Google 的 Chubby一个开源的实现，是 Hadoop 的分布式协调服务
- » 它包含一个简单的原语集，分布式应用程序可以基于它实现同步服务，配置维护和命名服务等
- » 它基于一套文件系统API实现，操作很小的数据节点`znodes`

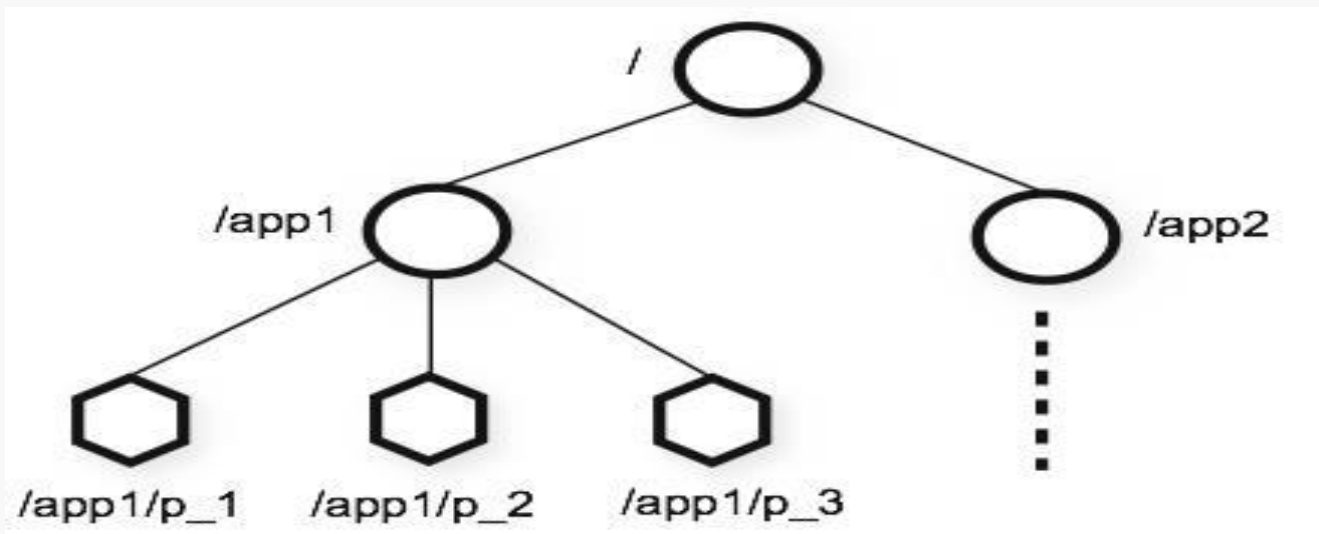


Zookeeper能帮我们做什么？

- » Hadoop,使用Zookeeper的事件处理确保整个集群只有一个NameNode,存储配置信息等.
- » HBase,使用Zookeeper的事件处理确保整个集群只有一个HMaster,察觉HRegionServer联机和宕机,存储访问控制列表等.



ZooKeeper数据模型



1. 分层命名空间。
2. 每个命名空间的节点都叫做“znode”。
3. 每个znode被路径区分（如：/app1）。
4. znode节点类型- 永久节点和临时节点。
5. 临时节点不能有子节点。
6. 每个znode节点有数据，也可以选择有子节点。
7. znode不能被重命名。
8. 可以给znode增加或者删除watchers。



观察 (watcher)

- » **Watcher** 在 **ZooKeeper** 是一个核心功能，**Watcher** 可以监控目录节点的数据变化以及子目录的变化，一旦这些状态发生变化，服务器就会通知所有设置在这个目录节点上的 **Watcher**，从而每个客户端都很快知道它所关注的目录节点的状态发生变化，而做出相应的反应
- » 可以设置观察的操作：`exists, getChildren, getData`
- » 可以触发观察的操作：`create, delete, setData`



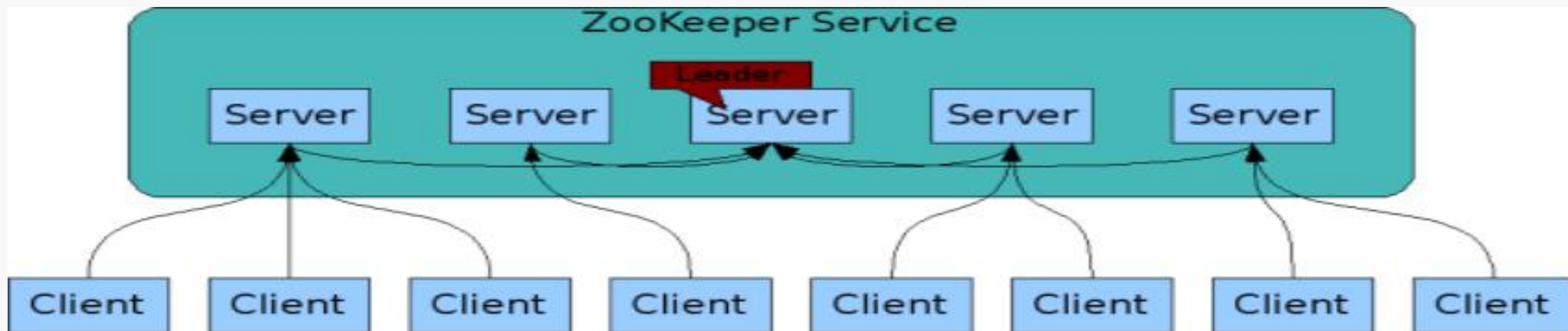
ACL (Access Control Lists)

- » 每个znode被创建时都会带有一个ACL列表，用于决定谁可以对它执行何种操作

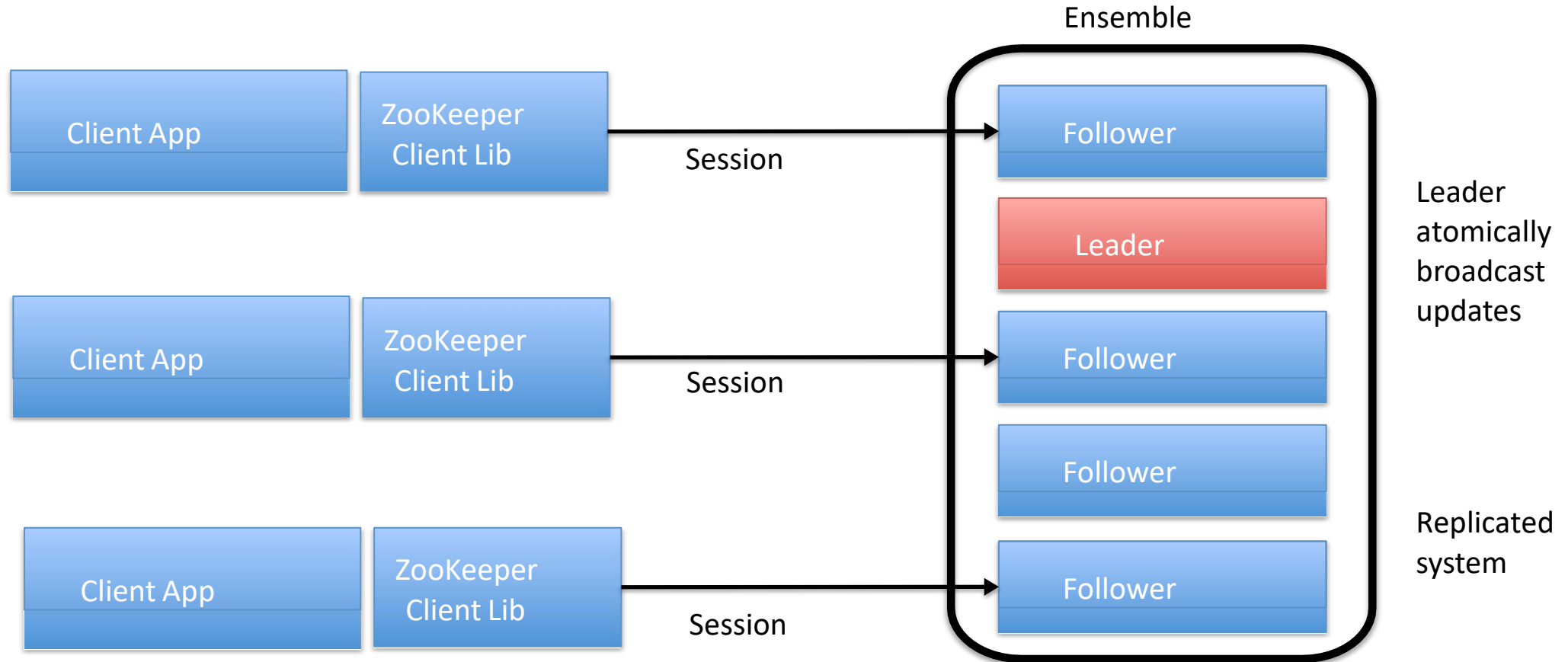
ACL权限	允许的操作
CREATE	create (子节点)
READ	getChildren getData
WRITE	setData
DELETE	delete (子节点)
ADMIN	setACL



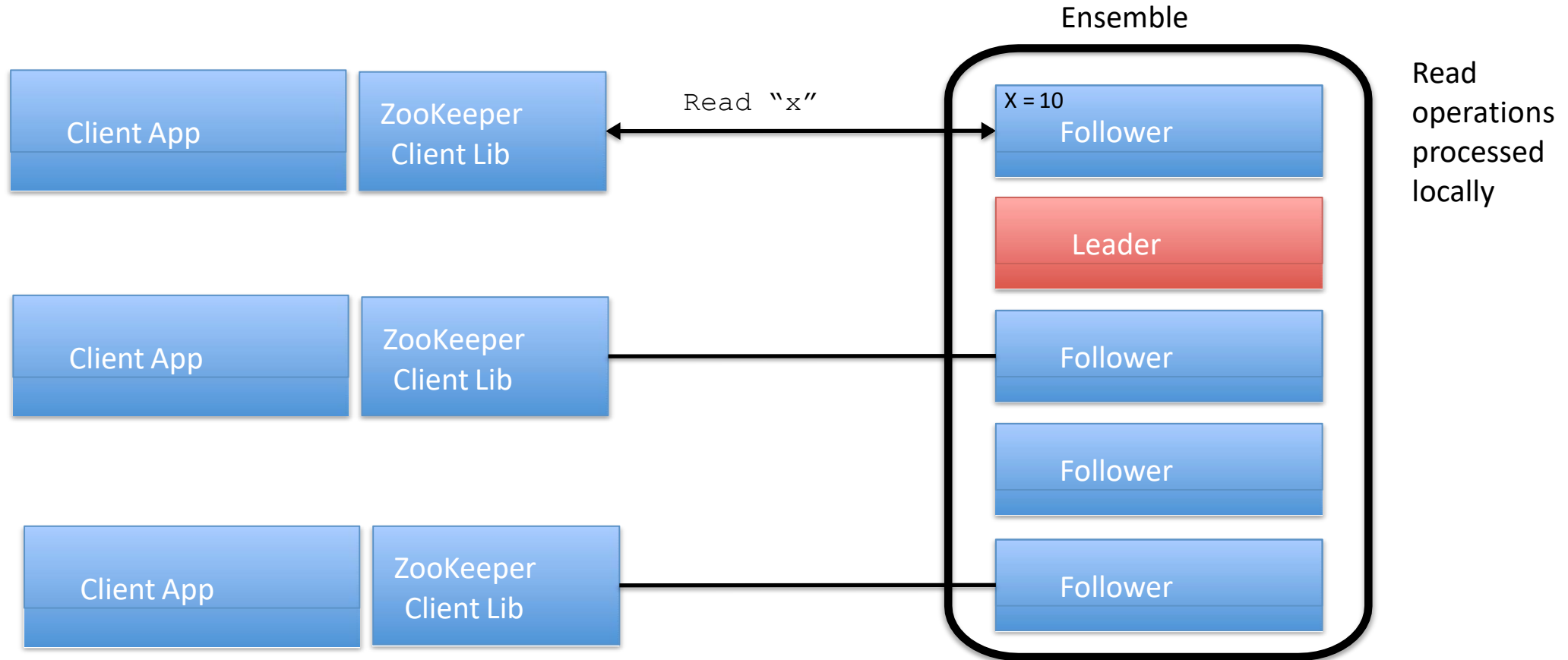
ZooKeeper服务架构--模型



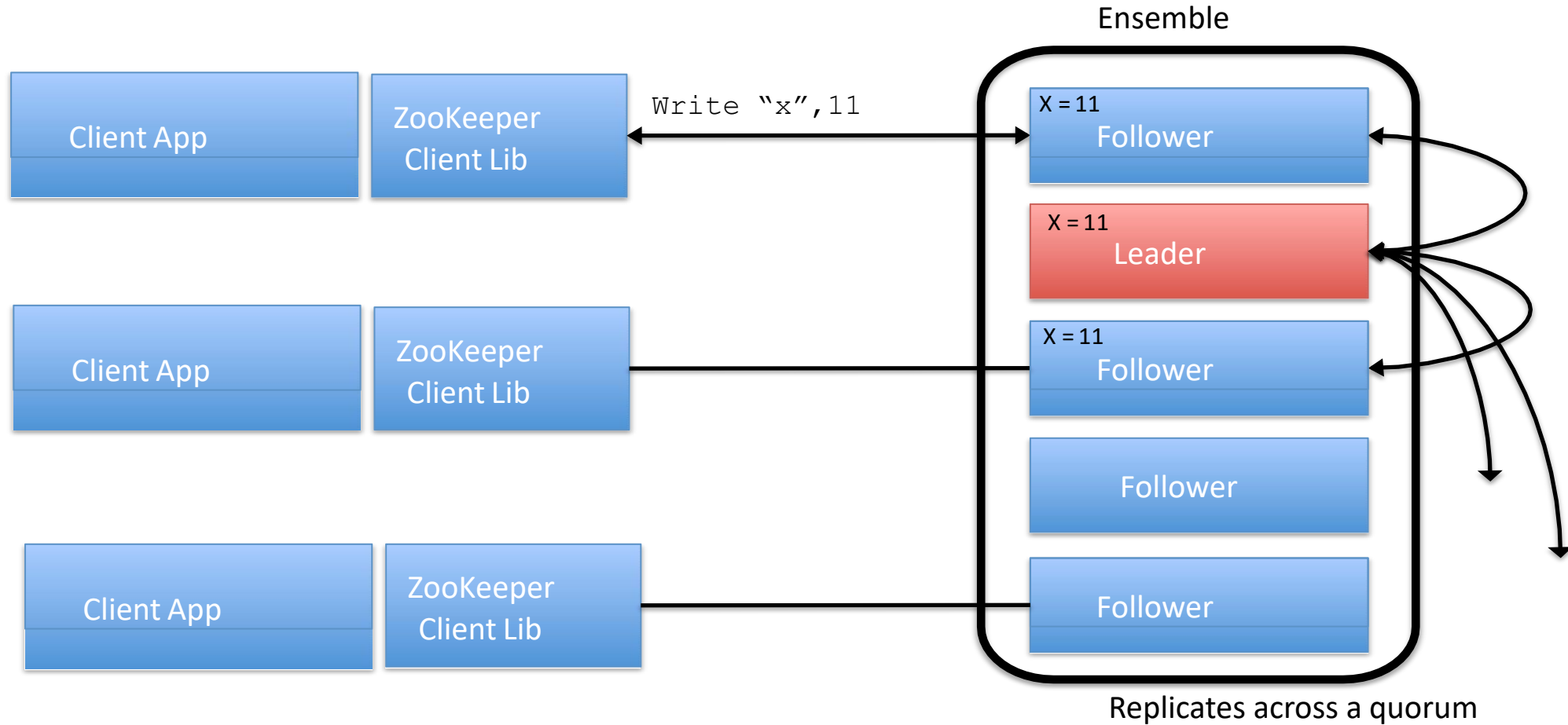
ZooKeeper: Overview



ZooKeeper: Read operations



ZooKeeper: Write operations





ZooKeeper的典型应用场景

1. 配置文件管理
2. 集群管理（client请求负载均衡）
3. 分布式协调/通知（例如：nimbus和supervisor）
4. 管理集群与Master选举
5. 分布式锁（集群唯一）
6. 分布式队列（多个客户访问同一组件）