

# ZooKeeper 集群分布式协 调服务

[www.huawei.com](http://www.huawei.com)





# 目标

- 学完本课程后，您将能够：
  - 掌握ZooKeeper概念
  - 掌握ZooKeeper系统架构
  - 掌握ZooKeeper关键特性



泰克教育  
TECH EDUCATION



# 目录

1. ZooKeeper简介
2. ZooKeeper在产品的位罝
3. 系统架构
4. 关键特性介绍
5. 与组件的关系

泰克教育  
TECH EDUCATION

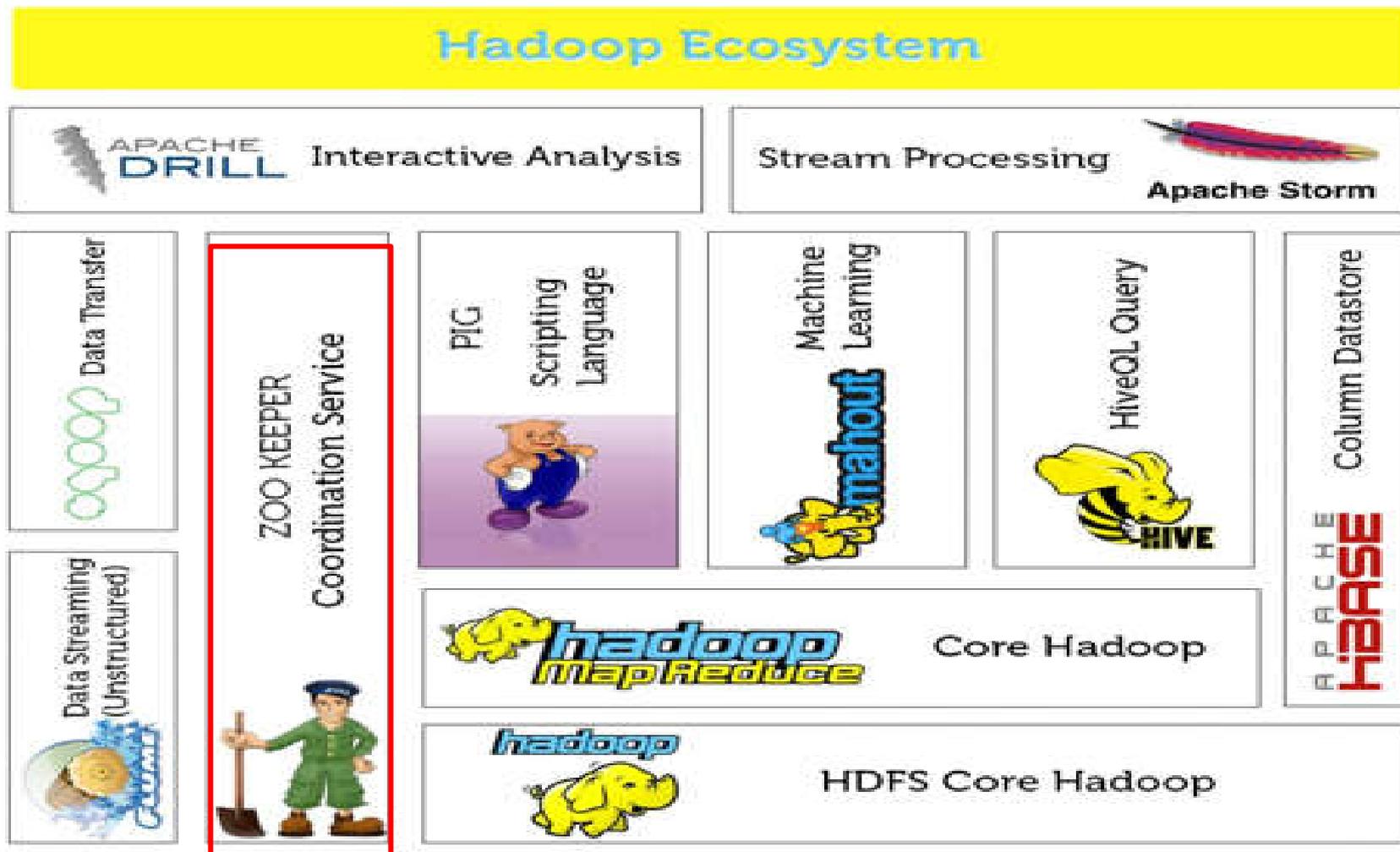
# ZooKeeper概述

- ZooKeeper 分布式服务框架主要是用来解决分布式应用中经常遇到的一些数据管理问题，提供分布式、高可用性的协调服务能力。
- 安全模式下ZooKeeper依赖于Kerberos和LdapServer进行安全认证，非安全模式则不依赖于Kerberos与Ldap。ZooKeeper作为底层组件广泛被上层组件使用并依赖，如Kafka，HDFS，HBase，Storm等。



东克教育  
TECH EDUCATION

# ZooKeeper简介



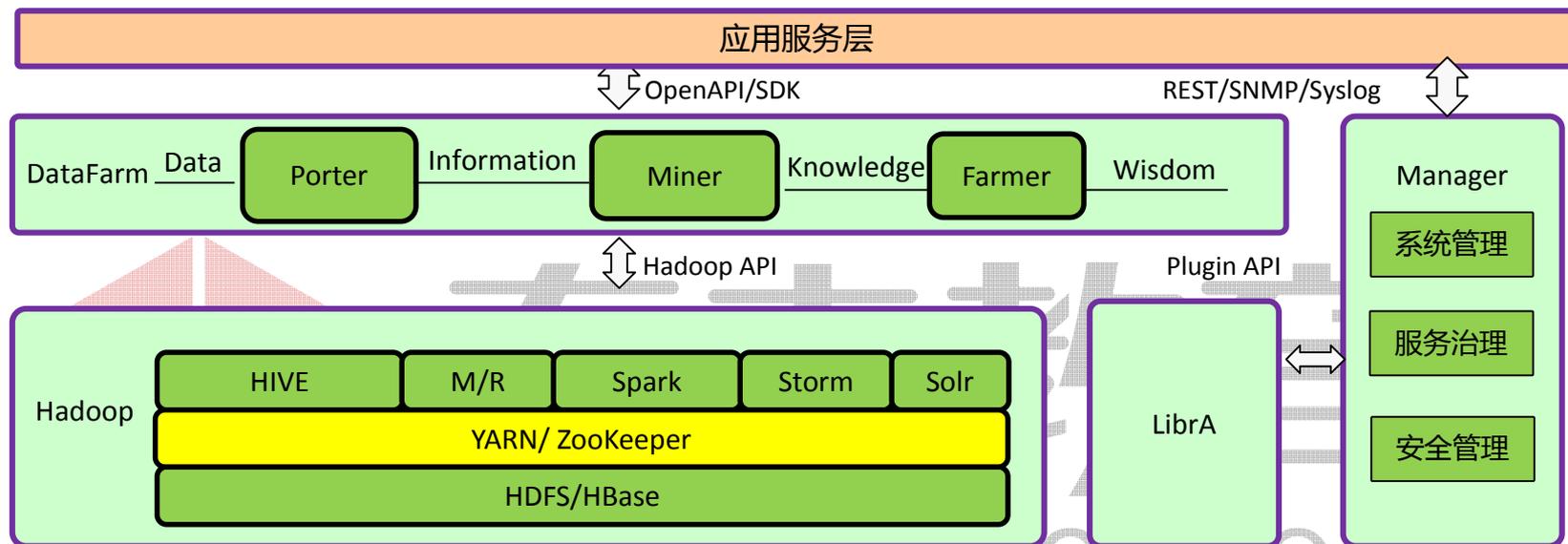


# 目录

1. ZooKeeper简介
- 2. ZooKeeper在产品的位罝**
3. 系统架构
4. 关键特性介绍
5. 与组件的关系

泰克教育  
TECH EDUCATION

# ZooKeeper在FusionInsight中的位置



ZooKeeper基于开源Apache ZooKeeper作为底层组件为上层组件提供服务,主要用于解决分布式应用中经常遇到的一些数据管理问题。

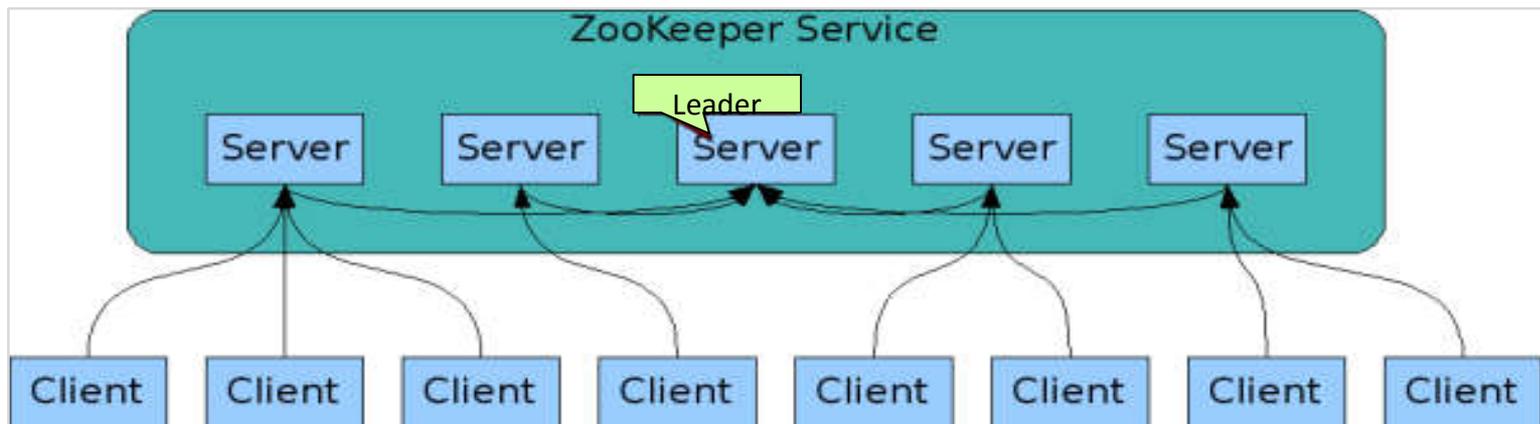


# 目录

1. ZooKeeper简介
2. ZooKeeper在产品的位罝
- 3. 系统架构**
4. 关键特性介绍
5. 与组件的关系

泰克教育  
TECH EDUCATION

# ZooKeeper服务架构 – 模型



- ZooKeeper集群由一组Server节点组成，这一组Server节点中只存在一个Leader的节点，其他节点都为Follower。
- 启动时选举出leader。
- ZooKeeper使用自定义的原子消息协议，保证了整个系统中的节点数据的一致性。
- Leader节点在接收到数据变更请求后，先写磁盘再写内存。

# ZooKeeper服务架构 - 容灾能力

- ZooKeeper能够完成选举即能够正常对外提供服务。
  - ZooKeeper选举时，当某一个实例获得了半数以上的票数时，则变为leader。
- 对于n个实例的服务，n可能为奇数或偶数。
  - n为奇数时，假定  $n = 2x + 1$ ，则成为leader的节点需获得x+1票，容灾能力为x。
  - n为偶数时，假定  $n = 2x + 2$ ，则成为leader的节点需要获得x+2票（大于一半），容灾能力为x。



# 目录

1. ZooKeeper简介
2. ZooKeeper在产品的位罝
3. 系统架构
4. **关键特性介绍**
5. 与组件的关系

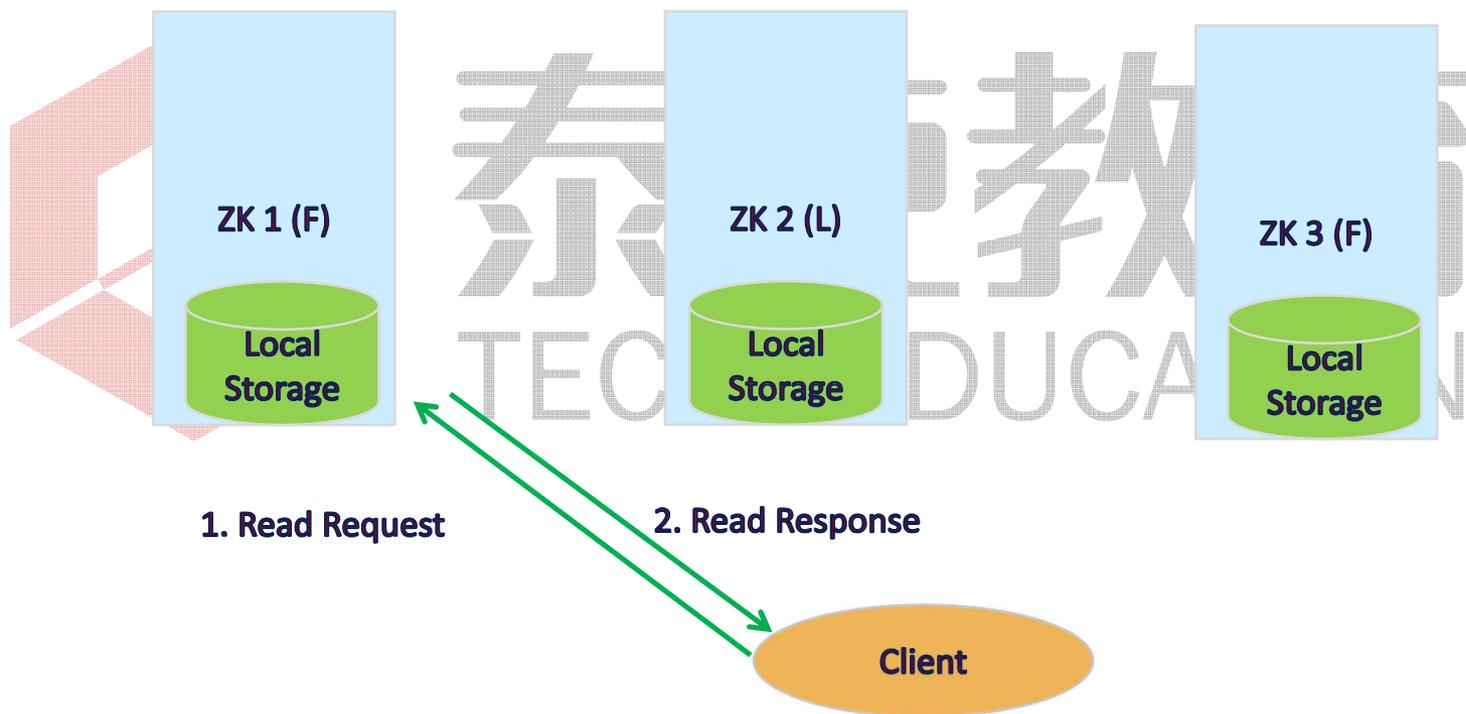
泰克教育  
TECH EDUCATION

# ZooKeeper 关键特性

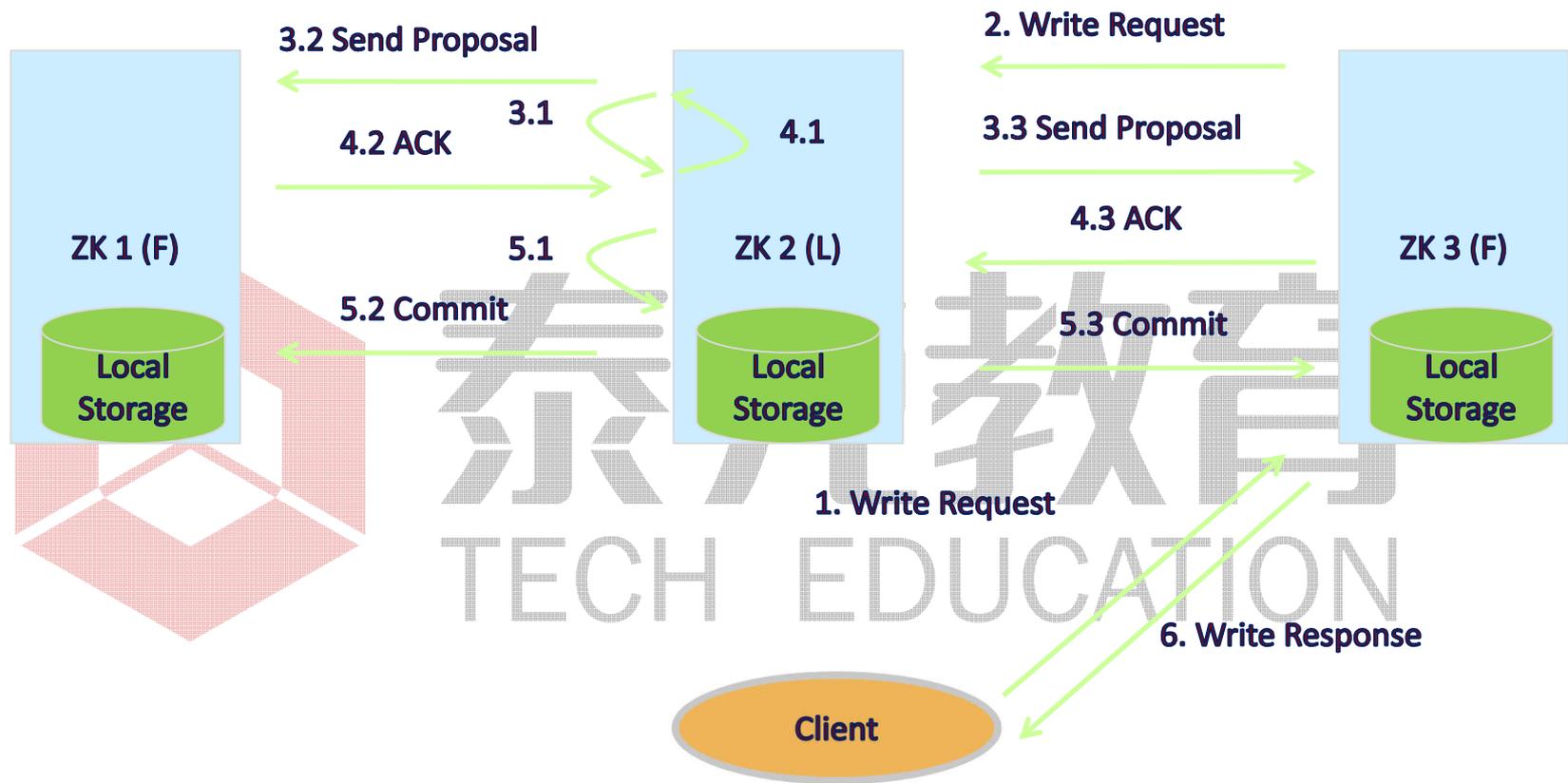
- 最终一致性：无论哪个server，对外展示的均是同一个视图。
- 实时性：保证客户端将在一个时间间隔范围内获得服务器的更新信息，或者服务器失效的信息。
- 可靠性：一条消息被一个server接收，它将被所有server接受。
- 等待无关性：慢的或者失效的client不会干预快速的client的请求，使得每个client都能有效的等待。
- 原子性：更新只能成功或者失败，没有中间状态。
- 顺序一致性：客户端所发送的更新会按照它们被发送的顺序进行应用。

# ZooKeeper读特性

- 由ZooKeeper的一致性可知，客户端无论连接哪个server，获取的均是同一个视图。所以，读操作可以在客户端与任意节点间完成。



# ZooKeeper写特性



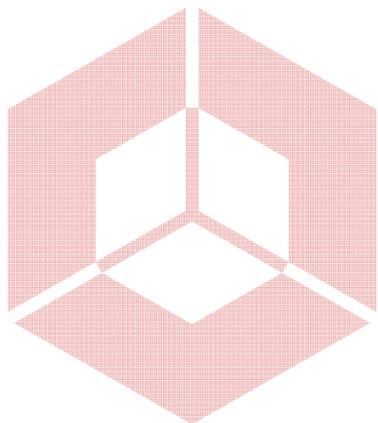
# ACL（访问控制列表）

ACL可以控制访问ZooKeeper的节点，只能应用于特定的znode上，而不能应用于该znode的所有子节点上。设置ACL命令为 `setAcl /znode scheme:id:perm`。

- Scheme为认证方式，ZooKeeper内置了4种方式：
  - world 一个单独的ID，表示任何人都可以访问。
  - auth 不使用ID，只有认证的用户可以访问
  - digest 使用username:password生成MD5哈希值作为认证ID。
  - IP 使用客户端主机IP地址来进行认证。
- Id：用来认证的字段，不同的scheme的认证方式不同。
- Perm：即permission，通过Acl认证的用户对该节点可拥有的操作权限。

# 日志增强

- Ephemeral node（临时节点）在session过期之后就会被系统删除。在审计日志中添加Ephemeral node被删除的审计日志，以便了解当时Ephemeral node的状态信息。



泰克教育  
TECH EDUCATION

# ZooKeeper客户端常用命令使用

- 调用ZooKeeper客户端，执行命令：

```
zkCli.sh -server 172.16.0.1:24002
```

- 创建节点：`create /node`
- 列出节点子节点：`ls /node`
- 创建节点数据：`set /node data`
- 获取节点数据：`get /node`
- 删除节点：`delete /node`
- 删除节点及所有子节点：`deleteall /node`

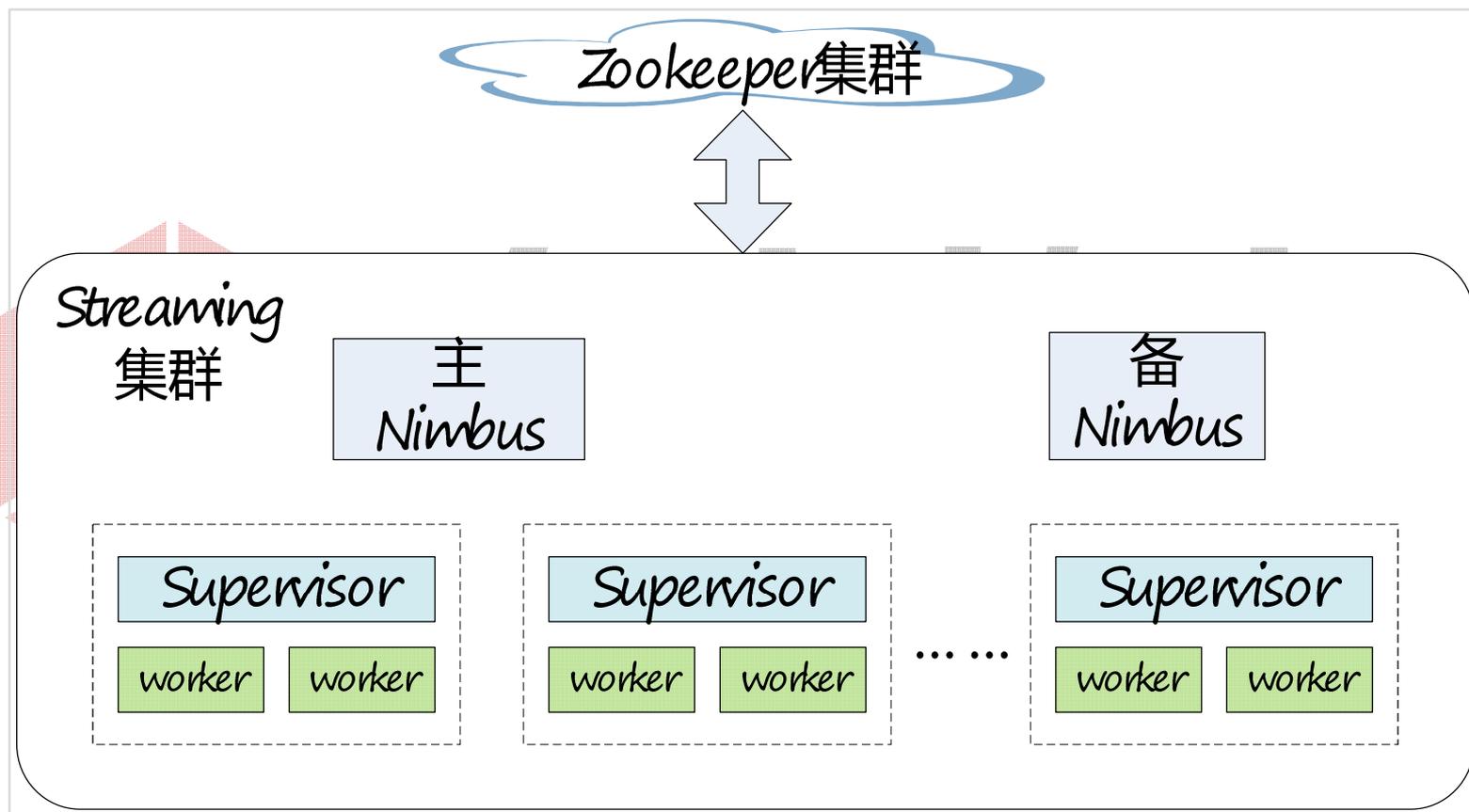


# 目录

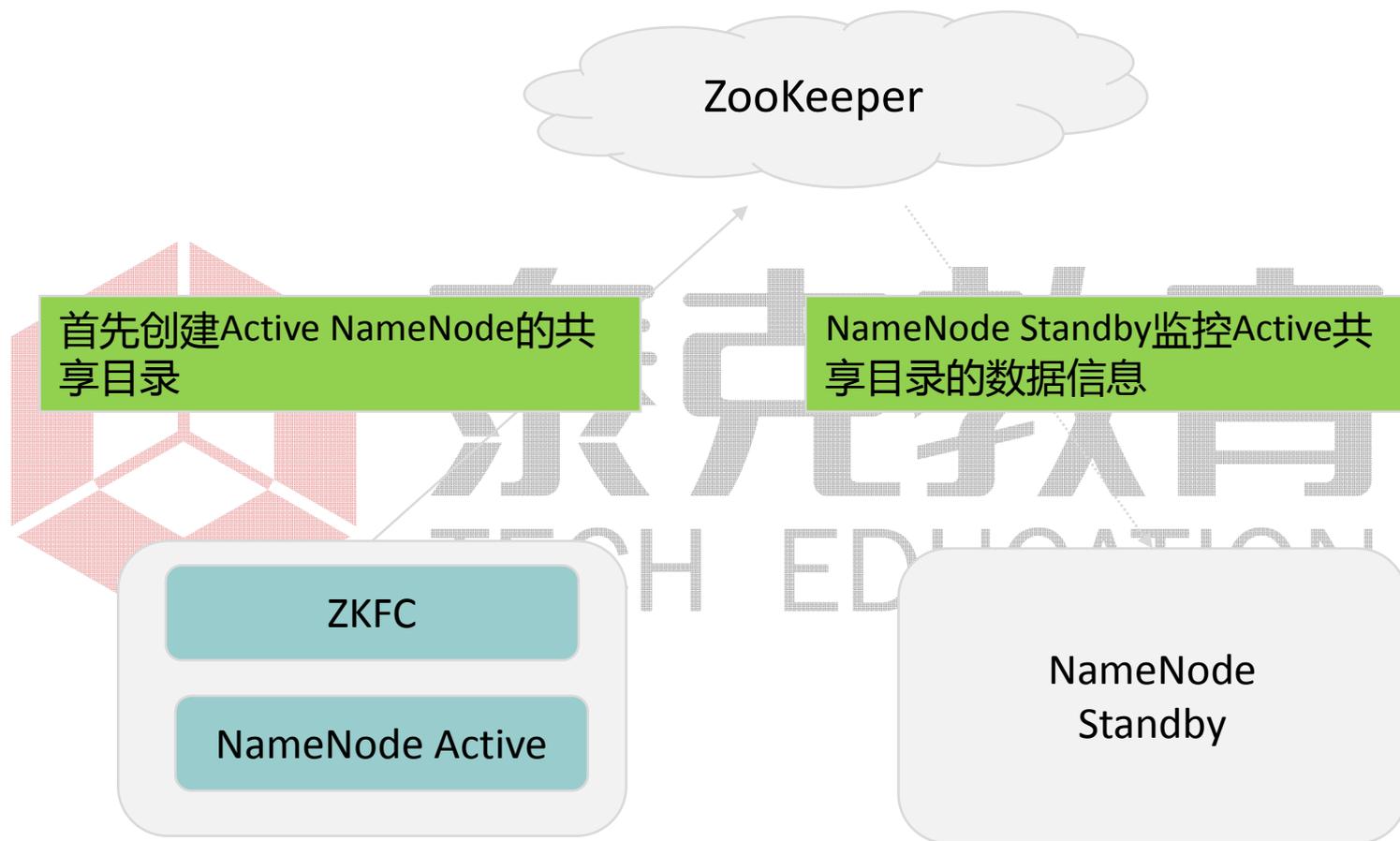
1. ZooKeeper简介
2. ZooKeeper在产品的位罝
3. 系统架构
4. 关键特性介绍
5. **与组件的关系**

泰克教育  
TECH EDUCATION

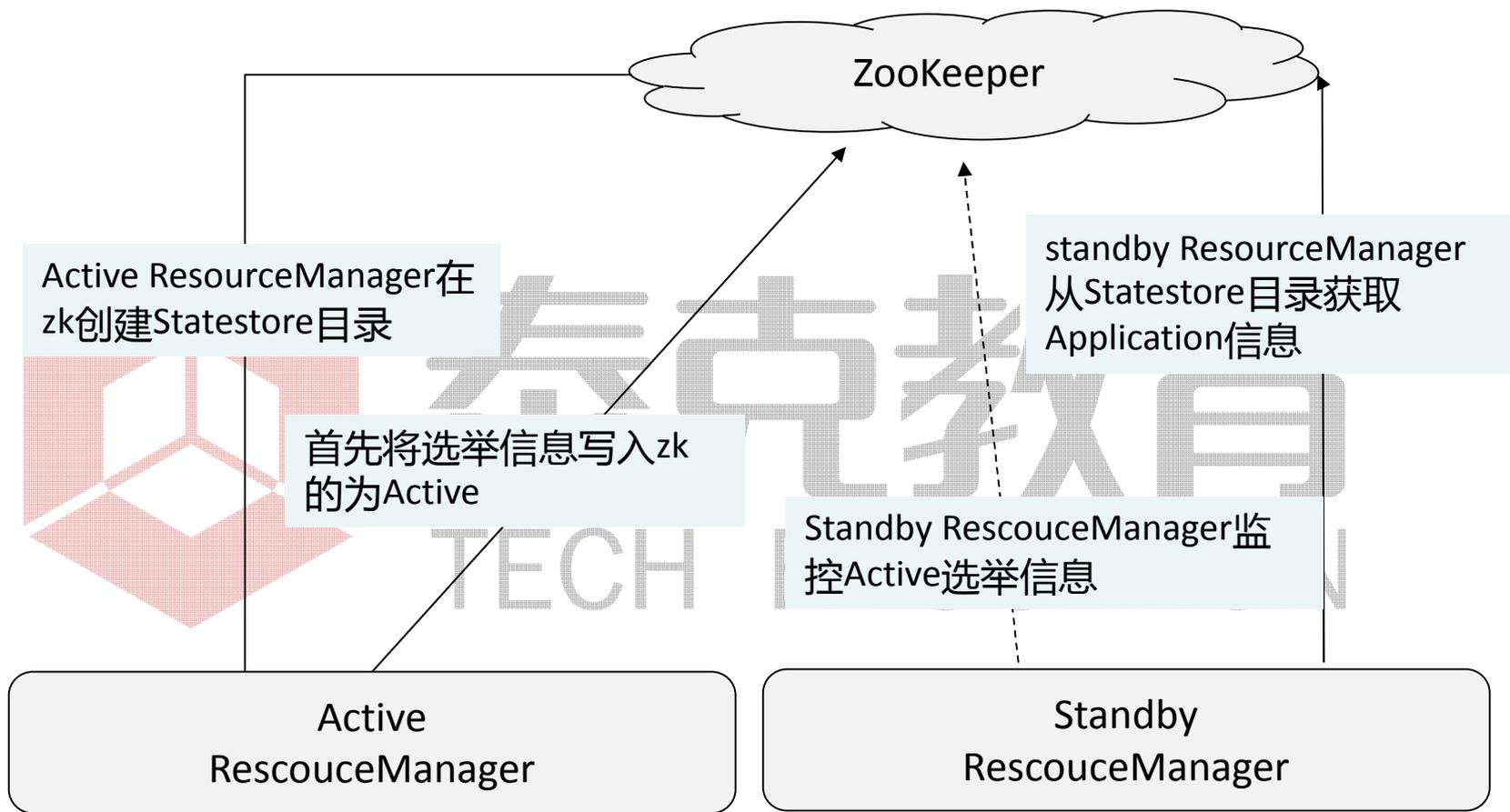
# ZooKeeper和Streaming



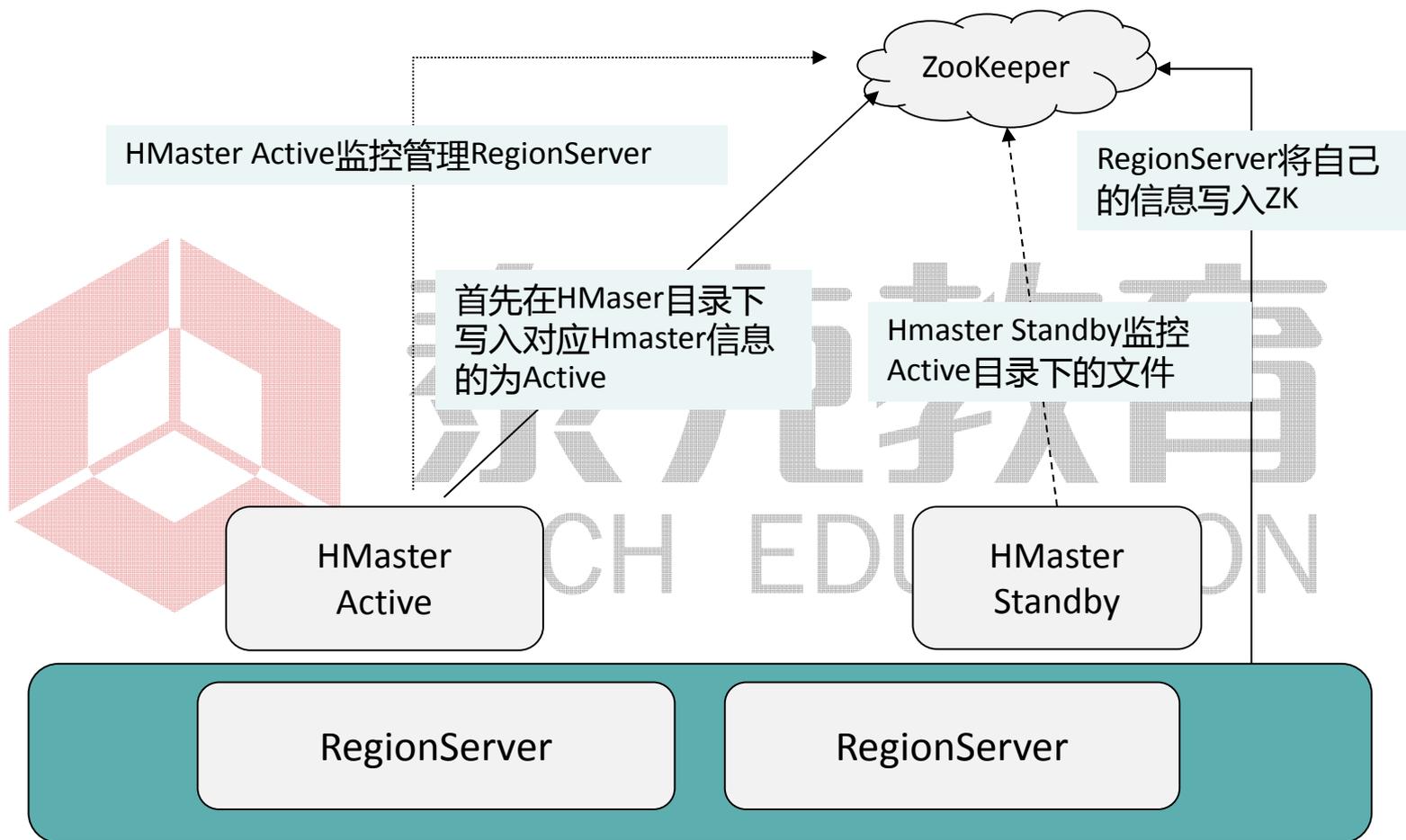
# ZooKeeper和HDFS



# ZooKeeper和YARN



# ZooKeeper和HBase



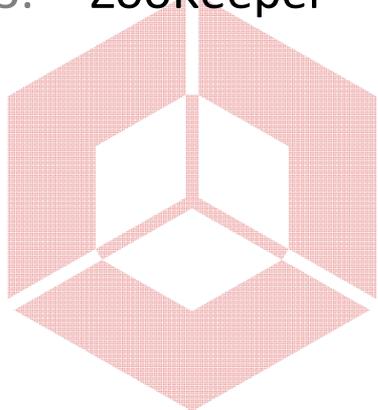


## 本章总结

- 本章节主要介绍知识点有：
  - ZooKeeper的作用以及在华为产品中的位置。
  - ZooKeeper的服务架构以及数据模型。
  - ZooKeeper的读写特性以及一致性的保证。
  - ZooKeeper节点的创建以及节点权限设定。
  - ZooKeeper与组件之间的关系。

## 思考题

1. ZooKeeper在集群中的位置及作用是什么？
2. ZooKeeper为什么建议奇数部署？
3. ZooKeeper一致性的含义是什么？

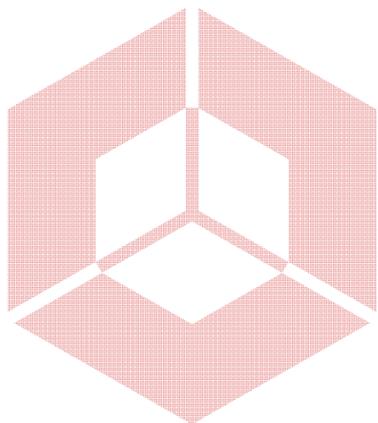


泰克教育  
TECH EDUCATION



## 更多信息

- 下载培训资料：
  - <http://support.huawei.com/learning/trainFaceDetailAction?lang=zh&pbiPath=term1000025185&courseId=Node1000009072>
- eLearning课程：
  - <http://support.huawei.com/learning/nodeQueryAction!loadTrainProjectInfo?lang=zh&pbiPath=term1000025185&courseId=Node1000009421&navId=MW000001>
- 考试大纲：
  - <http://support.huawei.com/learning/Certificate!toExamOutlineDetail?lang=zh&nodeId=Node1000003516>
- 模拟考试：
  - <http://support.huawei.com/learning/Certificate!toSimExamDetail?lang=zh&nodeId=Node1000004285>
- 认证流程：
  - [http://support.huawei.com/learning/NavigationAction!createNavi#navi\[id\]=\\_40](http://support.huawei.com/learning/NavigationAction!createNavi#navi[id]=_40)



谢谢

[www.huawei.com](http://www.huawei.com)

泰克教育  
TECH EDUCATION