# Using Paxos to Build a Scalable, Consistent,and Highly Available Datastore

## 中文翻译

## 英文原文

<https://pdos.csail.mit.edu/6.824/papers/spinnaker.pdf>

# 接口（API）

## get(key, colname, consistent)

### 简介

返回列值和版本号。

### 参数详解

#### Key：行的关键字

#### Colname：列名

#### Consistent：一致性级别，true表示强一致性（strong consistency），false表示时间轴一致性（timeline consistency）

## put(key, colname, colvalue)

## delete(key, colname)

## conditionalPut(key, colname, value, v)

只有在当前版本号等于v时，才可以成功操作，否则返回错误。

## conditionalDelete(key, colname, v)

# 架构

## 3份拷贝

3份拷贝放到相邻的3个节点中，如A、B、C。

## Cohort

一组节点，如3份拷贝所在的3个节点构成一个cohort。

## Zookeeper

### 分布式锁（distributed locks）

### Barriers

### Group membership

# 复制协议

## 两阶段

### Leader选举（leader election）

每个cohort中有一个Leader，其他的是Follower。

### 仲裁阶段（quorum phase）

## 时间

### 提交周期（commit period）

# 错误恢复

## 两个数据

### Last committed LSN（log sequence number）

### Last LSN in the log

## Follower恢复

### 本地恢复阶段（Local recovery phase）

### 追赶阶段（catch up phase）

## Leader接管

## Leader选举