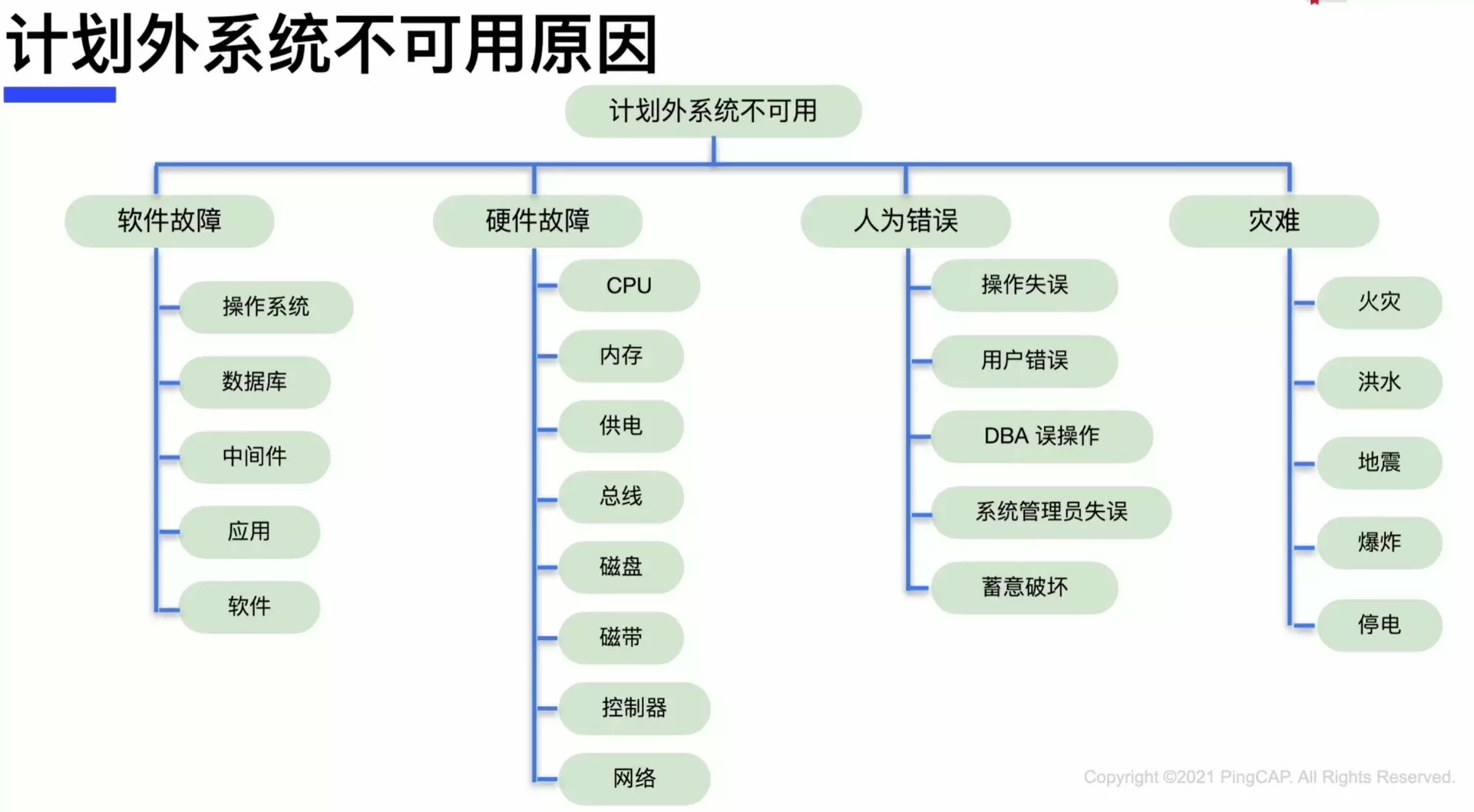


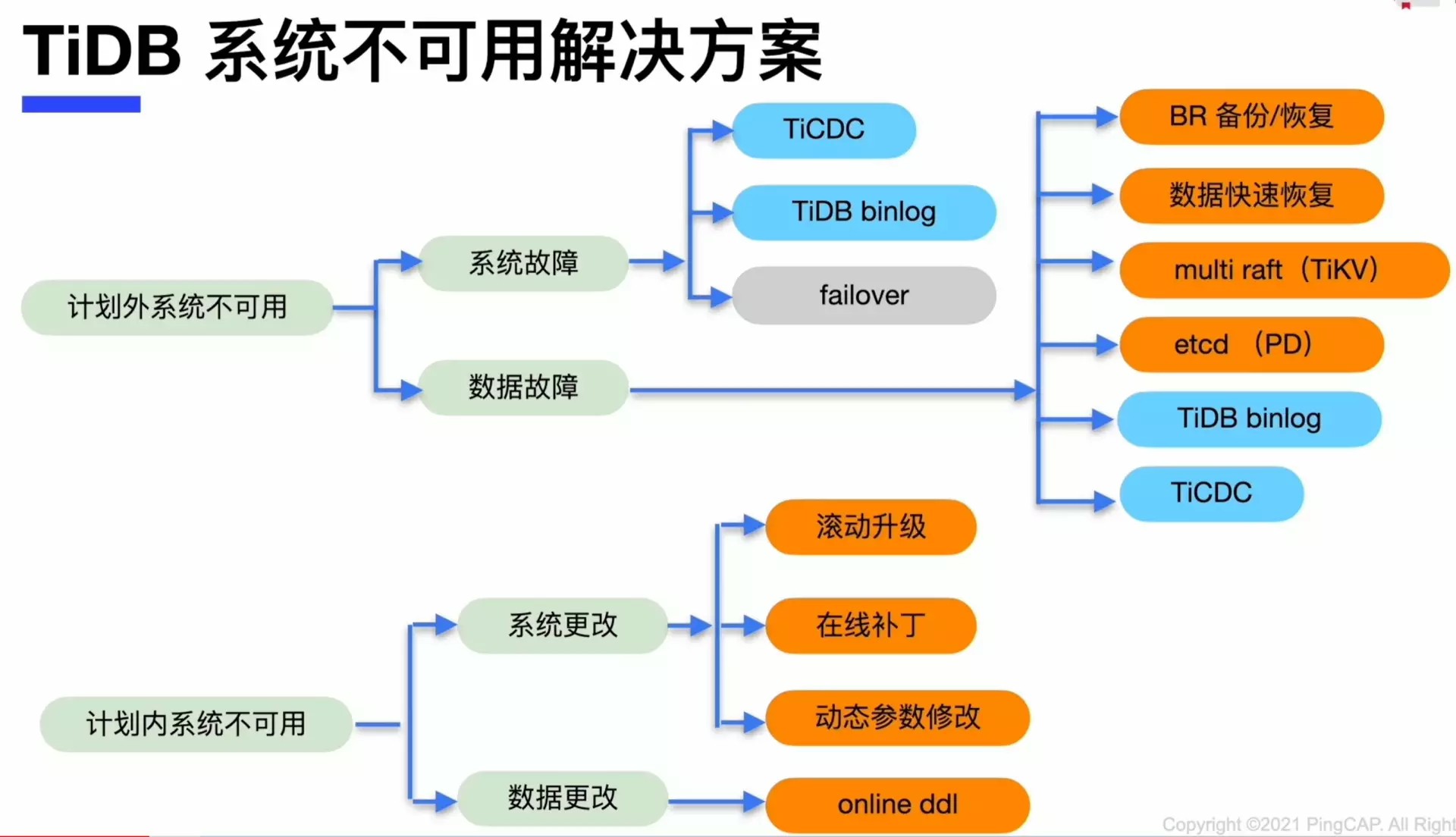
01、TiDB 数据库⾼可⽤概述

--2022-12-03 春雷

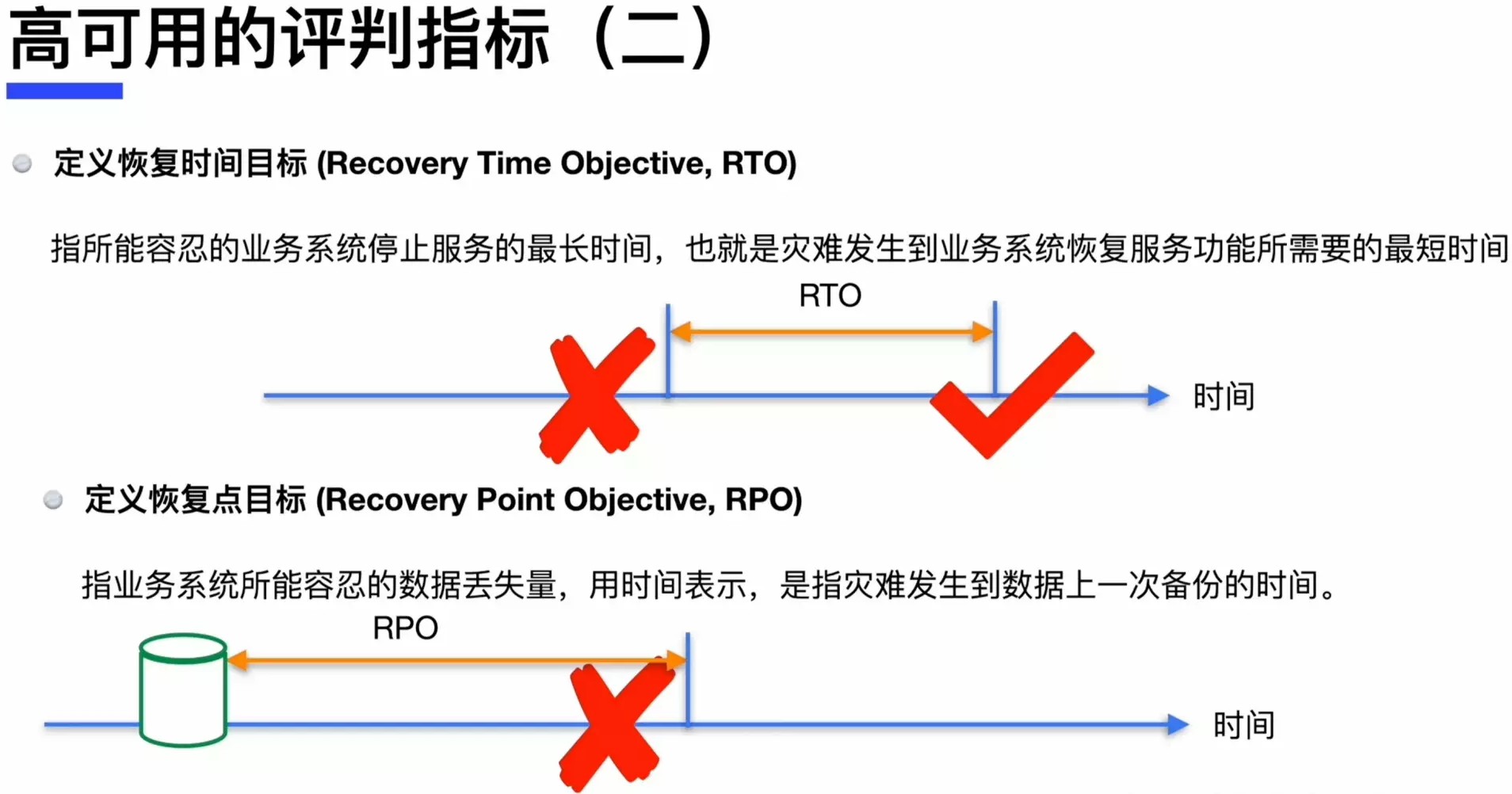
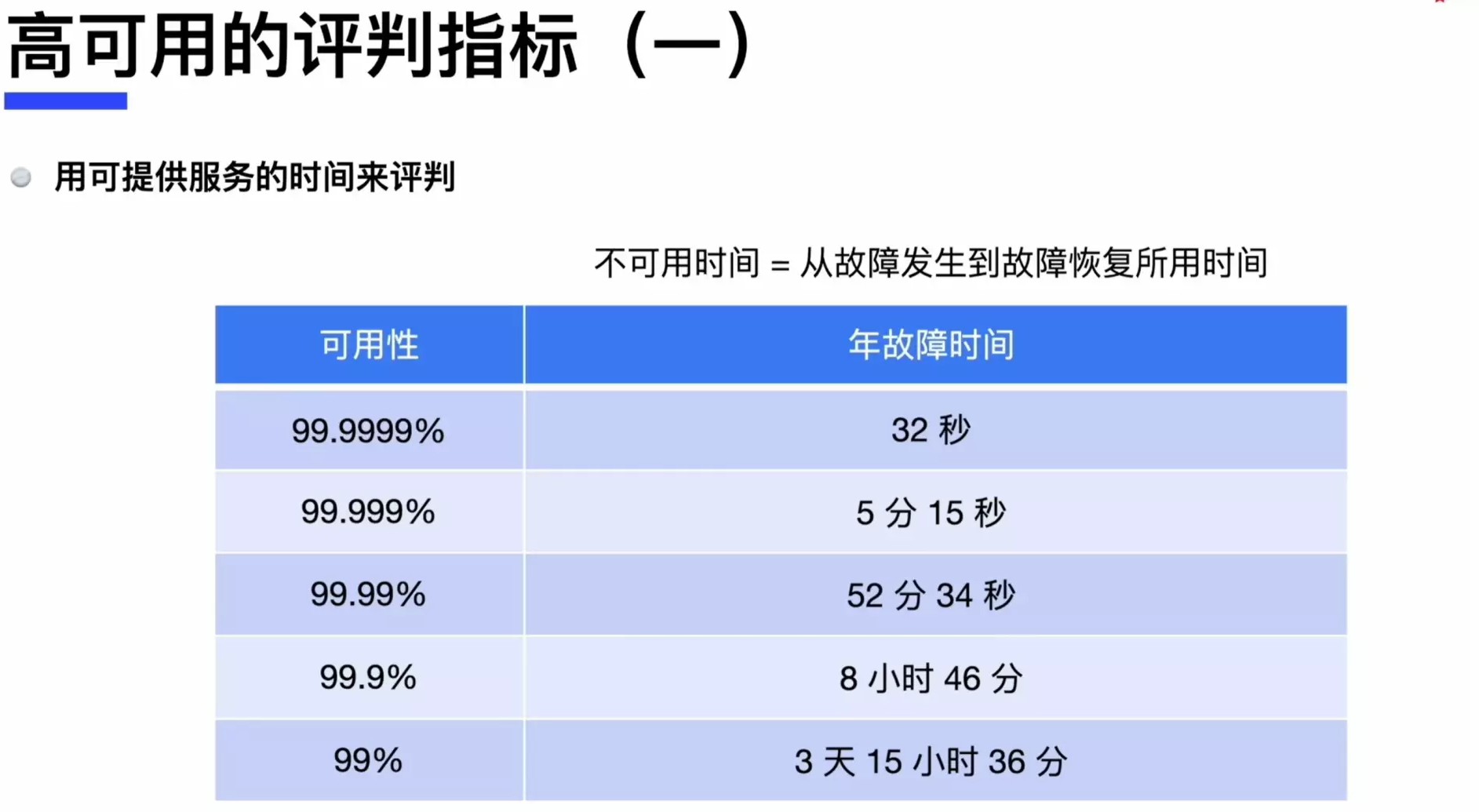
# 1、计划外系统不可⽤原因



2、TiDB系统不可⽤解决⽅案



# 3、⾼可⽤的评判指标

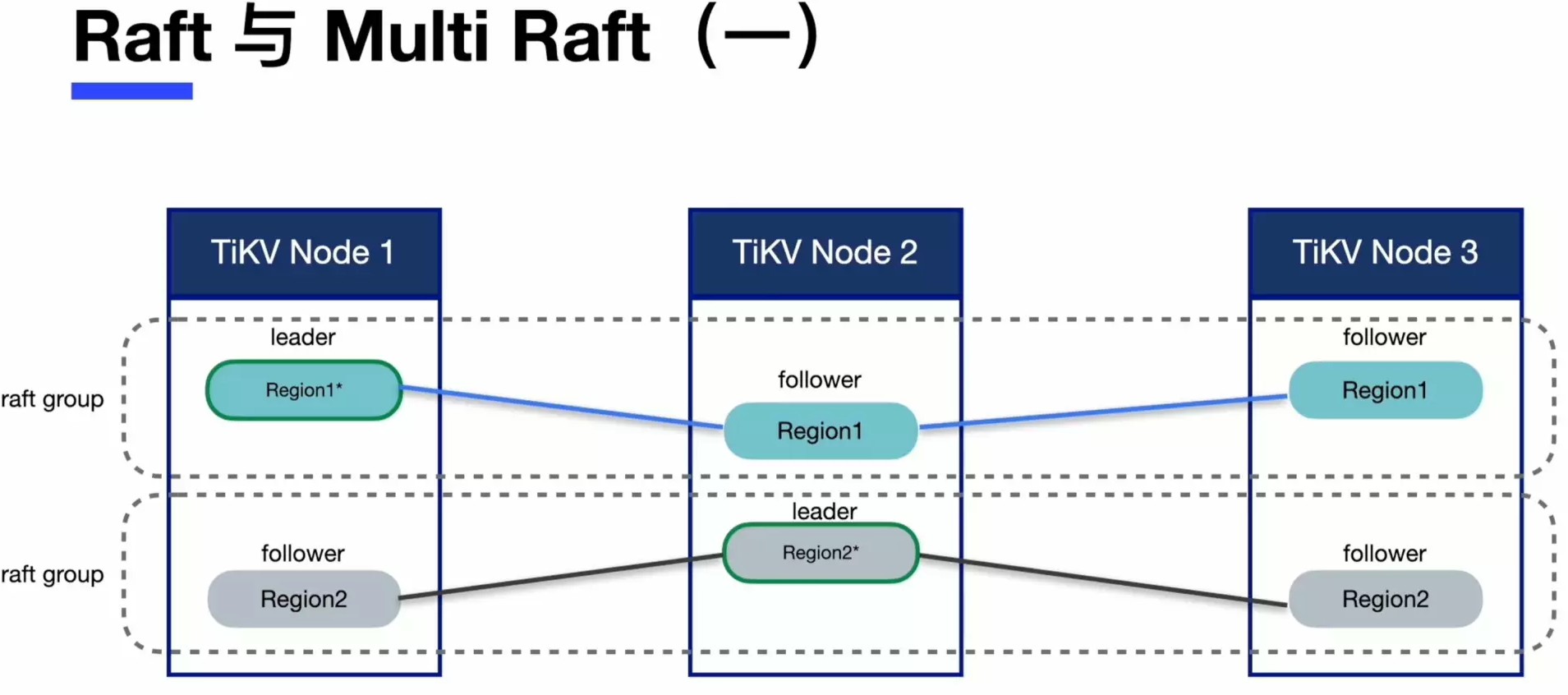




4、⾼可⽤的评判指标



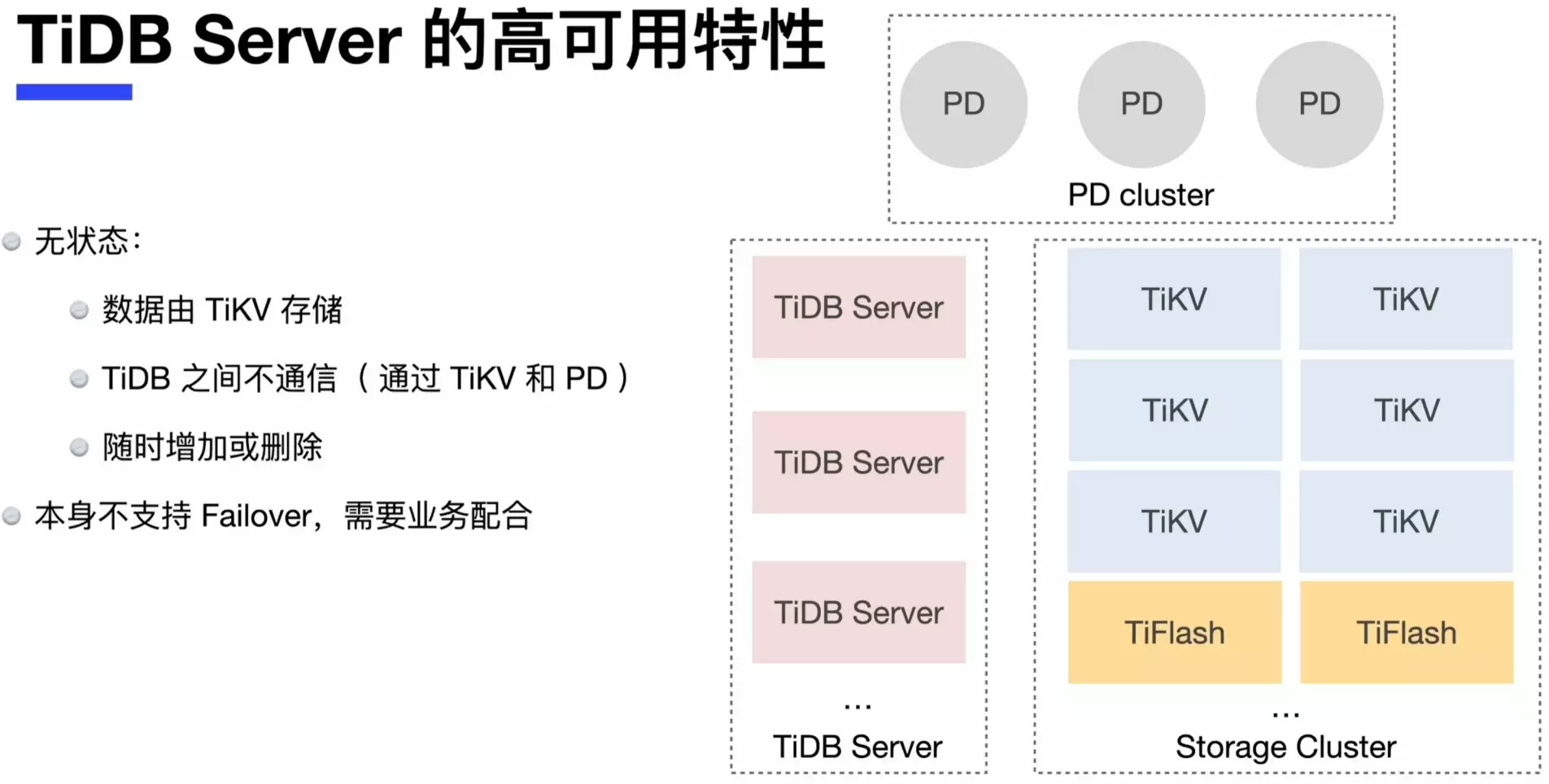
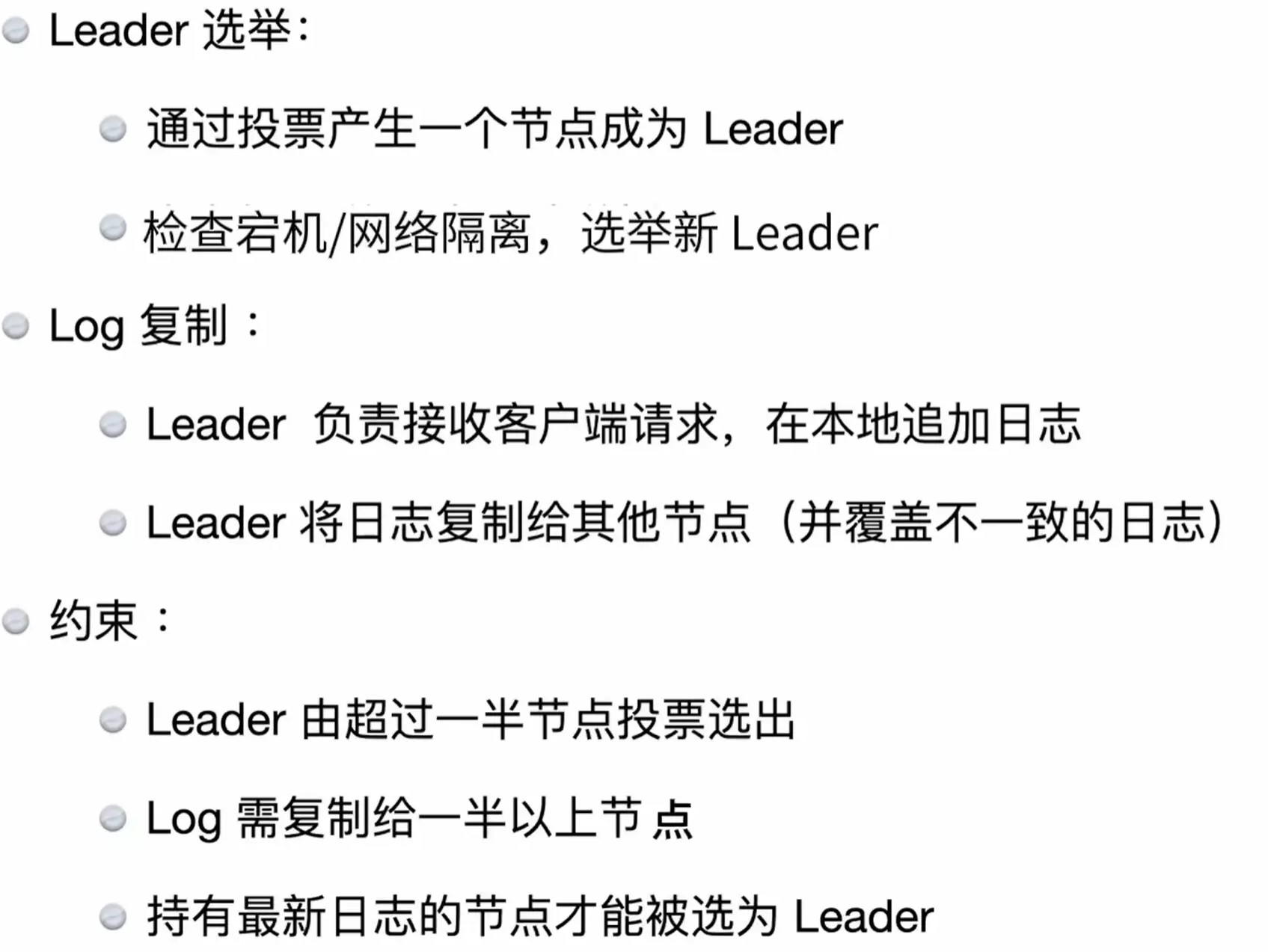
5、Raft 与 Multi Raft

Raft 是⼀种分布式⼀致性算法，在 TiDB 集群的多种组件中，PD 和 TiKV 都通过 Raft实现了数据的容灾。Raft 的灾难恢复能⼒通过如下机制实现：

Raft 成员的本质是⽇志复制和状态机。Raft 成员之间通过复制⽇志来实现数据同 步；Raft 成员在不同条件下切换⾃⼰的成员状态，其⽬标是选出 leader 以提供对外服务。

Raft 是⼀个表决系统，它遵循多数派协议，在⼀个 Raft Group 中，某成员获得⼤多数投票，它的成员状态就会转变为 leader。也就是说，当⼀个 Raft Group 还保有⼤多数节点 (majority) 时，它就能够选出 leader 以提供对外服务。

TiKV以region 为单位存储最⼩副本：3个

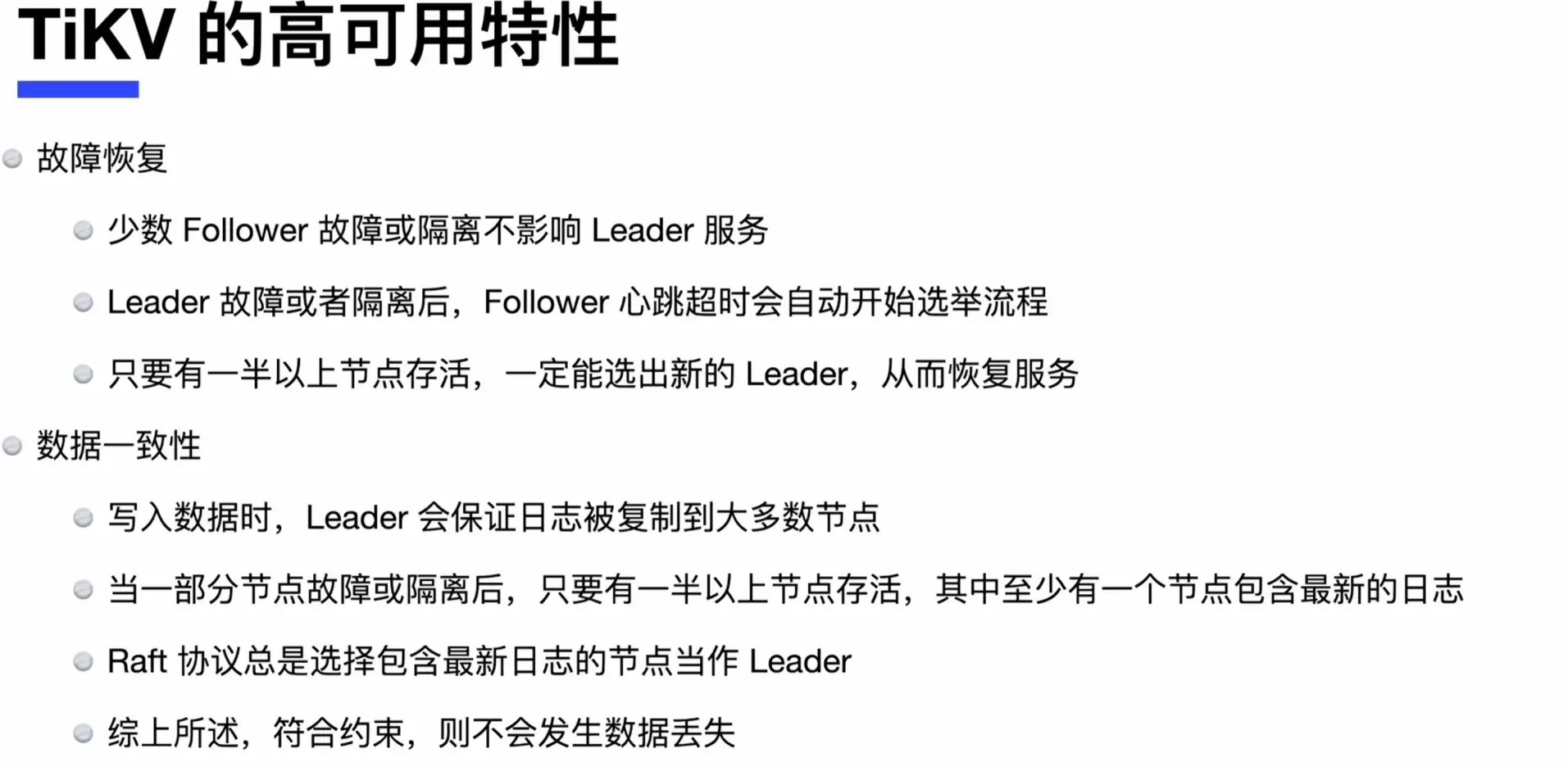
多个region 组成raft group多个raft group 为Multi Raft



6、TiDB Server的⾼可⽤特性



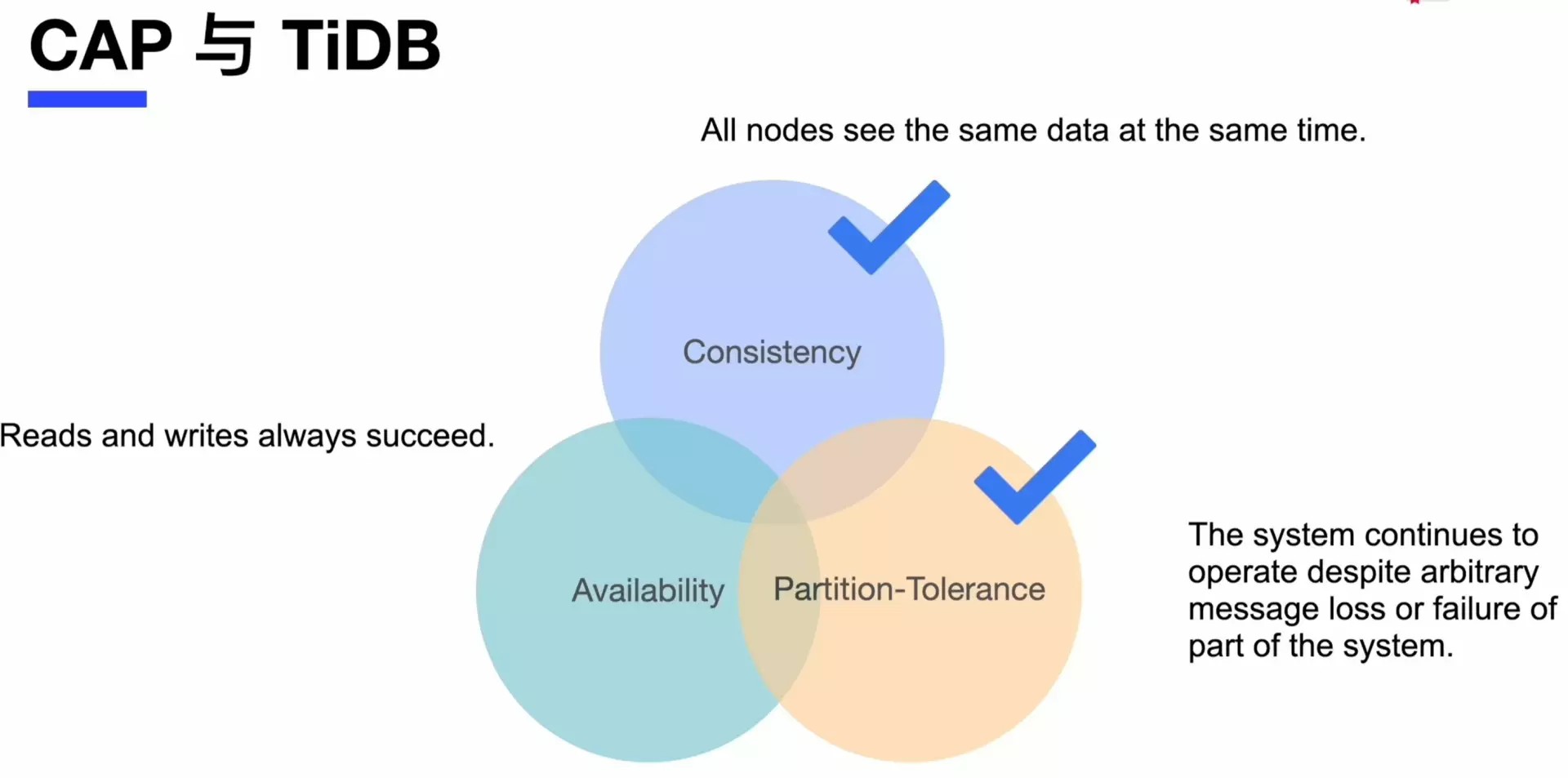
7、TiKV的⾼可⽤特性





8、PD的⾼可⽤特性

9、CAP与TiDB

【分布式数据库的CAP】:



（1）C（Consistency）：强⼀致性

“all nodes see the same data at the same time”,即更新操作成功并返回客户端后，所有节点在同⼀时间的数据完全⼀致，这就是分布式的⼀致性。⼀致性的问题在并发系统中不可避免，对

于客户端来说，⼀致性指的是并发访问时更新过的数据如何获取的问题。从服务端来看，则是更新如何复制分布到整个系统，以保证数据最终⼀致。



（2）A（Availability）：⾼可⽤性

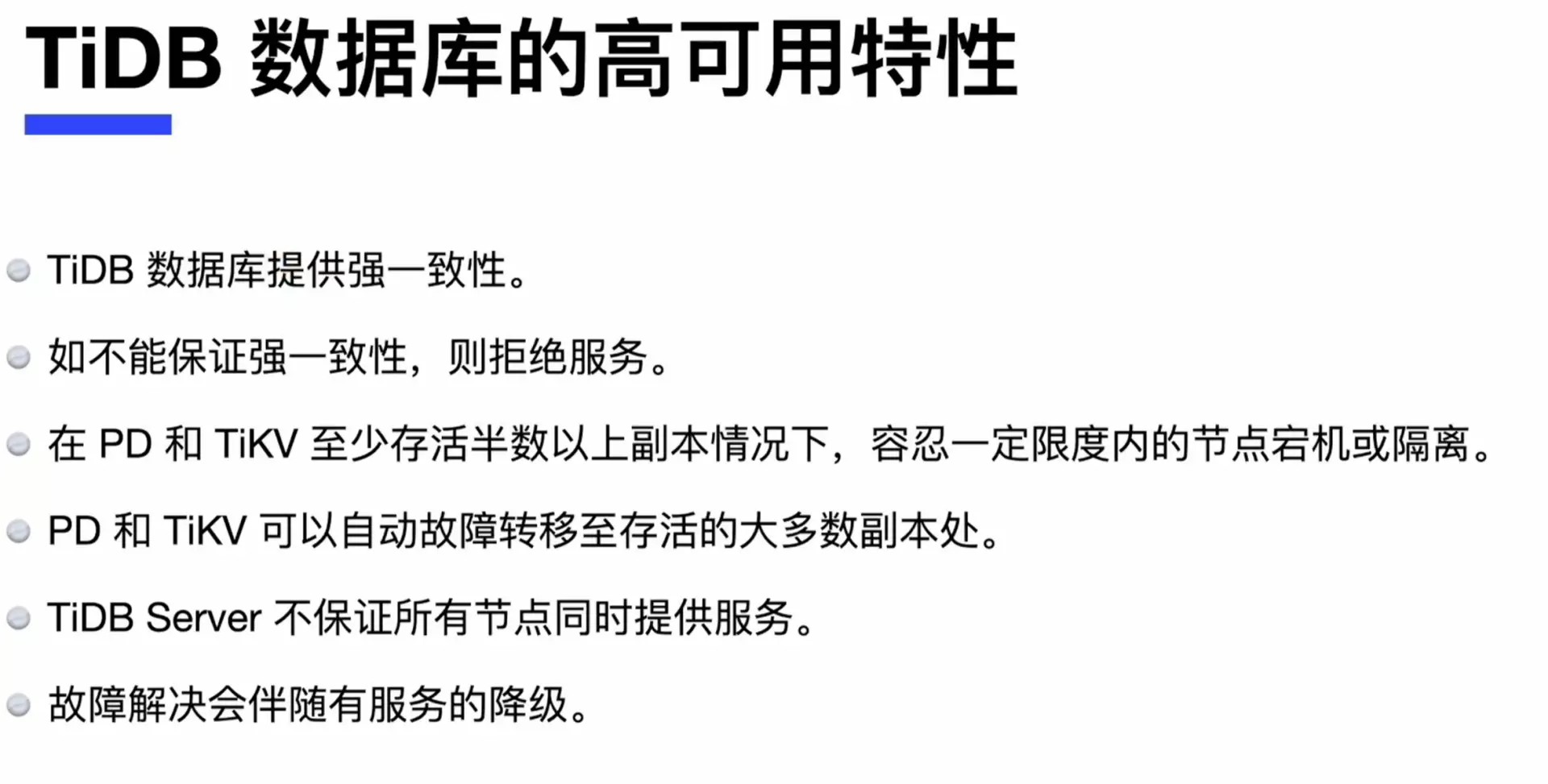
可⽤性指“Reads and writes always succeed”，即服务⼀直可⽤，⽽且要是正常的响应时间。好的可⽤性主要是指系统能够很好的为⽤户服务，不出现⽤户操作失败或者访问超时等⽤户体验不好的情况。



（3）P（Partition tolerance）：分区容错性

即分布式系统在遇到某节点或⽹络分区故障时，仍然能够对外提供满⾜⼀致性或可⽤性的服务。分区容错性要求能够使应⽤虽然是⼀个分布式系统，⽽看上去却好像是在⼀个可以运转正常的整体。⽐如现在的分布式系统中有某⼀个或者⼏个机器宕掉了，其他剩下的机器还能够正常运转满

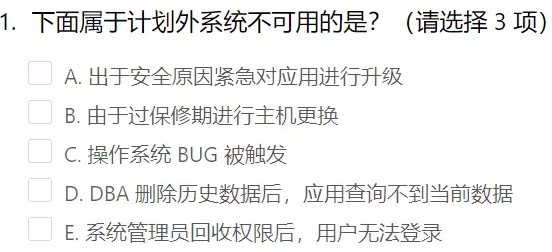
⾜系统需求，对于⽤户⽽⾔并没有什么体验上的影响。 TiDB是CP：即保证⼀致性与分区容错性





10、TiDB数据库的⾼可⽤特性

注：只要在⼀个节点写⼊成功了，只要其他节点可⽤，就可以查询到

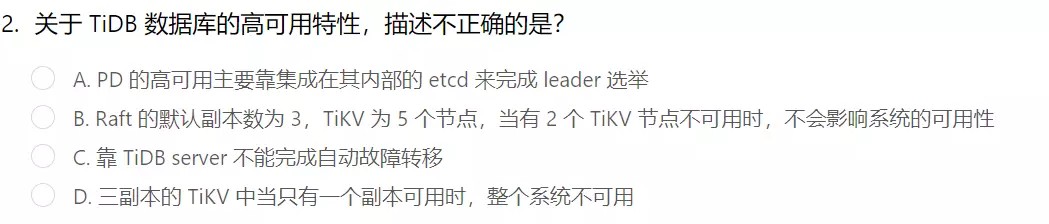


11、⼩测试

答案：CDE解析：

A：已知的升级，计划内 B：已知的更换，计划内 C：BUG是计划外 D：误操作删除，计划外

E：误操作回收了业务在⽤的账号权限，计划外



答案：B解析：

当宕的2个tikv节点上，包含的是⼀个raft group的2个region，则会有影响