3.6 拜占庭容错算法

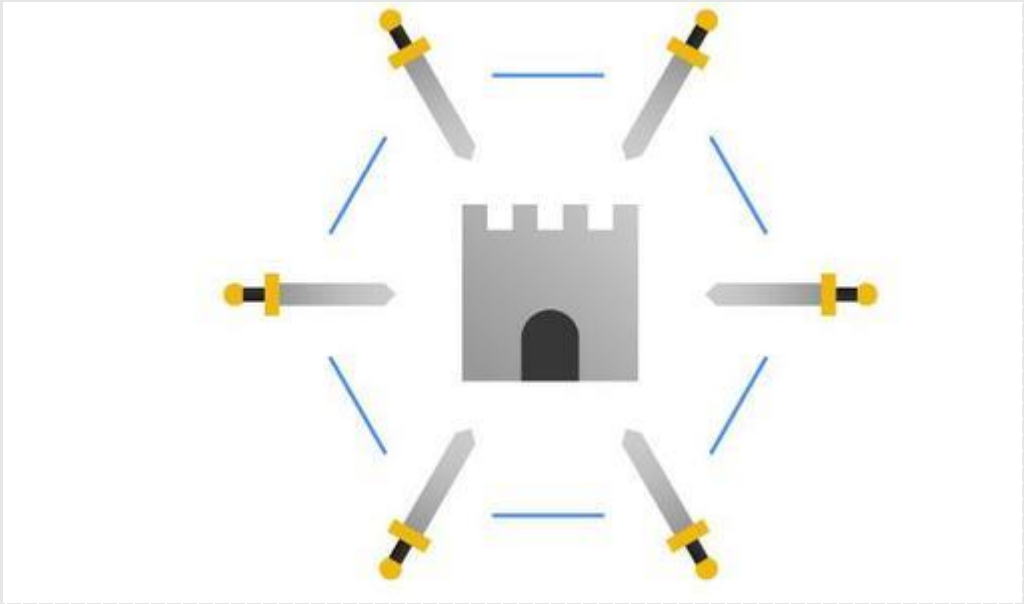
01 拜占庭容错共识算法

拜占庭容错，英文全称Byzantine Fault Tolerance，简称BFT。

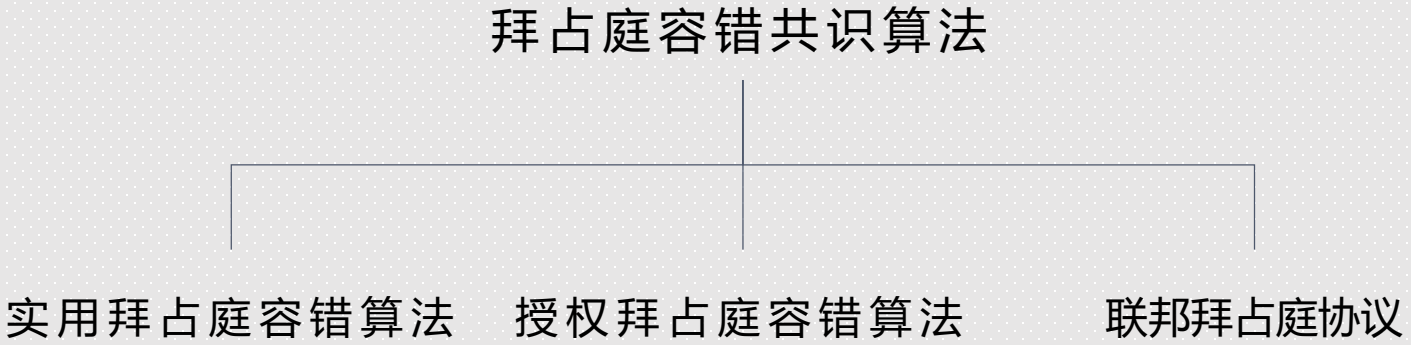
拜占庭错误：叛徒，也就是恶意节点，它为了阻挠真实信息的传递以及有效一致的达成，会向各个节点发送前后不一致的信息。

能够处理拜占庭错误的这种容错性，就叫做拜占庭容错。

拜占庭容错共识算法，就是假设区块链网络环境包括运行正常的服务器、故障的服务器和破坏者的服务器情况下，如何在正常的节点间形成对网络状态的共识。



02 拜占庭容错共识算法的种类



这些算法在具体的实现形式上有所不同，但都有速度快、支持高并发、可扩展的特点，通常被用于私有链或者联盟链。

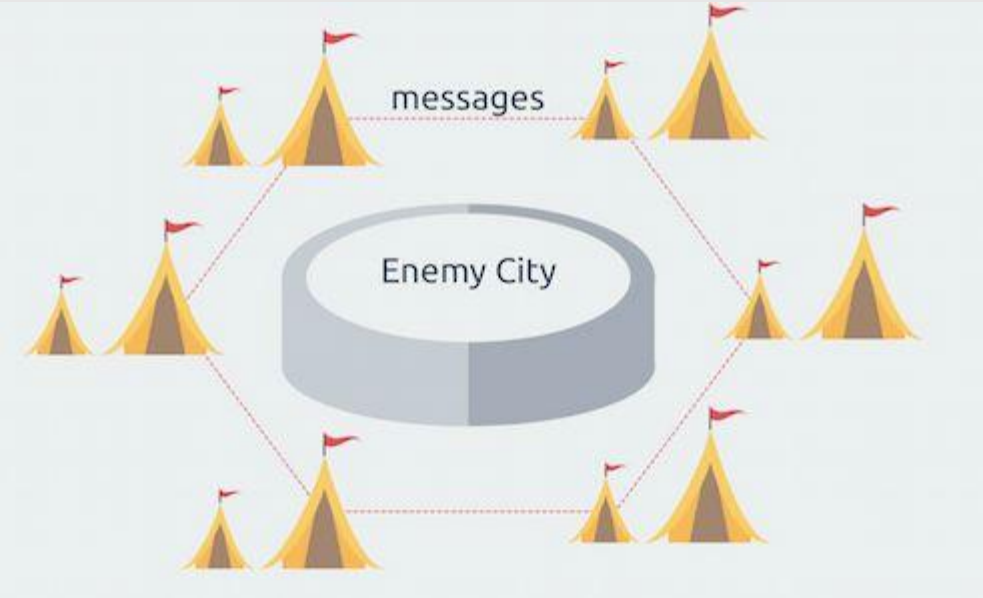
03 实用拜占庭容错算法

实用拜占庭容错算法，英文全称，Practical Byzantine Fault Tolerance，简称PBFT。

该算法是Miguel Castro (卡斯特罗)和Barbara Liskov（利斯科夫）在1999年提出来的，解决了原始拜占庭容错算法效率不高的问题，将算法复杂度由指数级降低到多项式级，使得拜占庭容错算法在实际系统应用中变得可行。

目前，实用拜占庭容错算法已经得到了比较广泛的使用，比如，在超级账本Fabric0.6中主要使用的就是这种算法。它可以在失效节点不超过总节点数1/3的情况下保证消息传递的正确可靠。

04 PBFT算法工作流程



战争开始了，分散在不同地方的将军们之间互相传递信息。每一个收到命令的将军都要去询问其他人，他们收到的命令是什么。

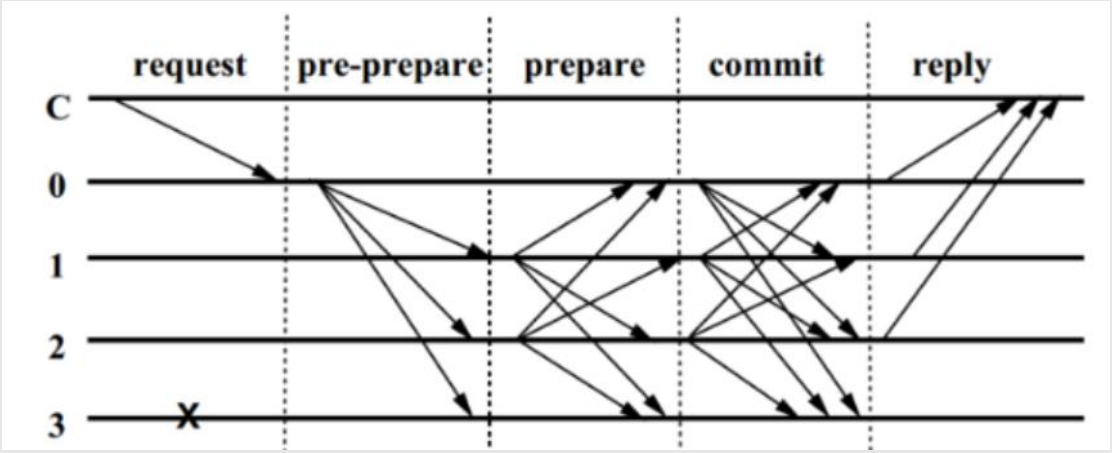
PBFT，本质上就是利用通信次数换取可靠性。

每个命令的执行都需要节点间两两交互去核验消息，这产生了比较高的通信代价。

05 PBFT算法工作流程

PBFT是一种状态机副本复制算法，即服务作为状态机进行建模，状态机在分布式系统的不同节点进行副本复制。**每个状态机的副本都保存了服务的状态，同时也实现了服务的操作。**所有的副本在一个被称为视图的轮换过程中运作。在某个视图中，一个副本作为主节点，其它的副本节点作为备份节点。主节点通过随机算法选出，用来负责与提案的客户端通信。主节点选出后，客户端发送提案给主节点，主节点将客户端请求进行编号，然后发送预准备消息给所有的副本节点。

06 PBFT算法工作流程

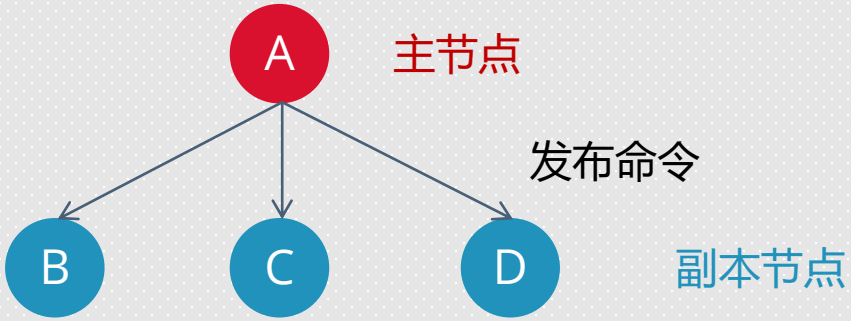


（C是客户端，0是主节点，123是副本节点其中3是无效节点）

每一个副本节点在收到来自主节点的预准备消息之后，都要检查消息的正确性，然后发送准备消息给除了自己以外的其他所有副本节点。同时它也会收到其他副本节点发来的准备消息。在收到消息后，副本节点对其他节点的准备消息进行验证，如果正确就将准备消息写入消息日志，集齐规定数量的准备消息之后，它就进入准备状态。

副本节点进入准备状态后，在全网范围内广播commit消息，当副本节点集齐规定数量个验证过的commit消息后，就表示请求处理完毕，当前网络中的大部分节点已经达成共识，于是发送处理结果给客户端，运行客户端的请求.

07 PBFT举例

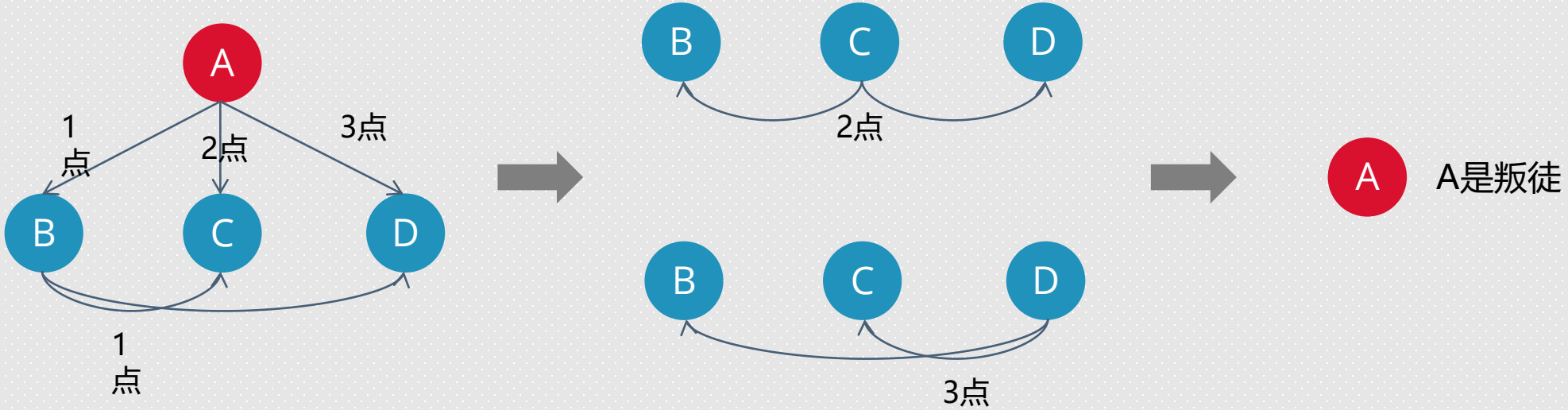


两种可能：

① A是叛徒；

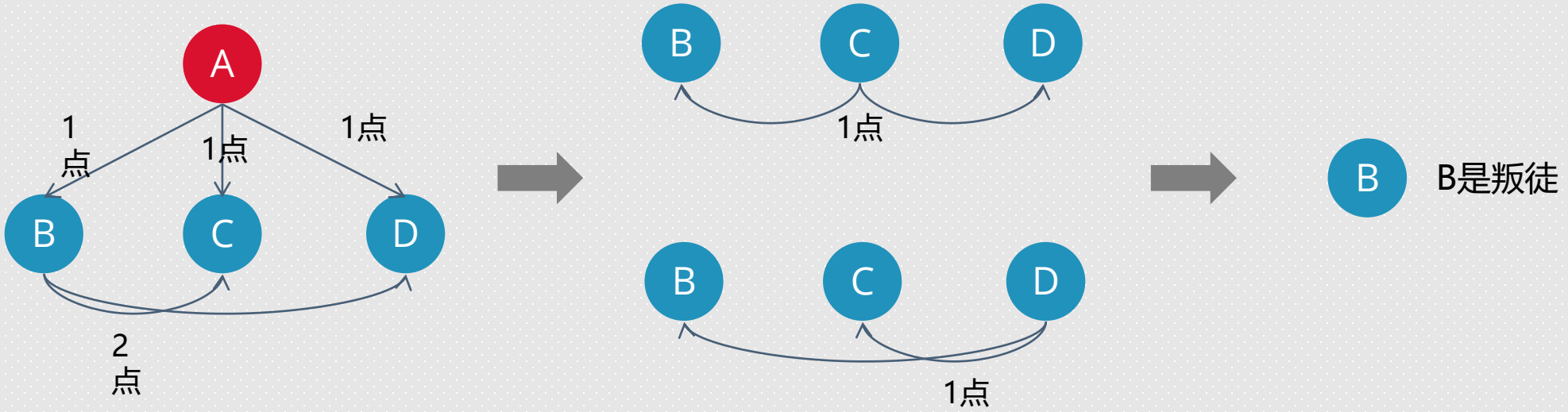
② B、C、D中存在叛徒。

① A是叛徒



08 PBFT举例

② B是叛徒



当节点数大于等于4个的时候，1个无效节点的存在并不会影响消息的传递。推广来说，当存在n个无效节点时，只要总节点数超过3n个，消息传递的正确性就能得到保证，这也是拜占庭算法的容错率。