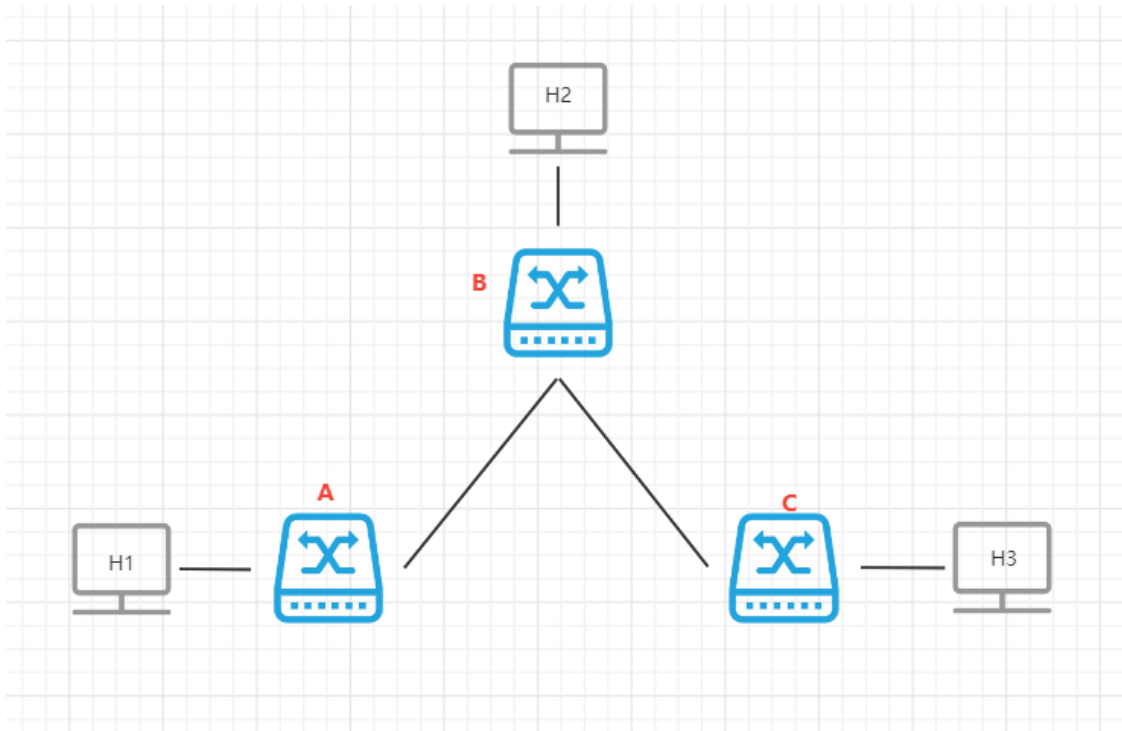


本篇文章来认识下以太网交换机所使用的生成树协议，我们来看看为什么要使用此协议，我们的基于交换机的以太网到底遇到了什么问题？走，让我们即刻出发！

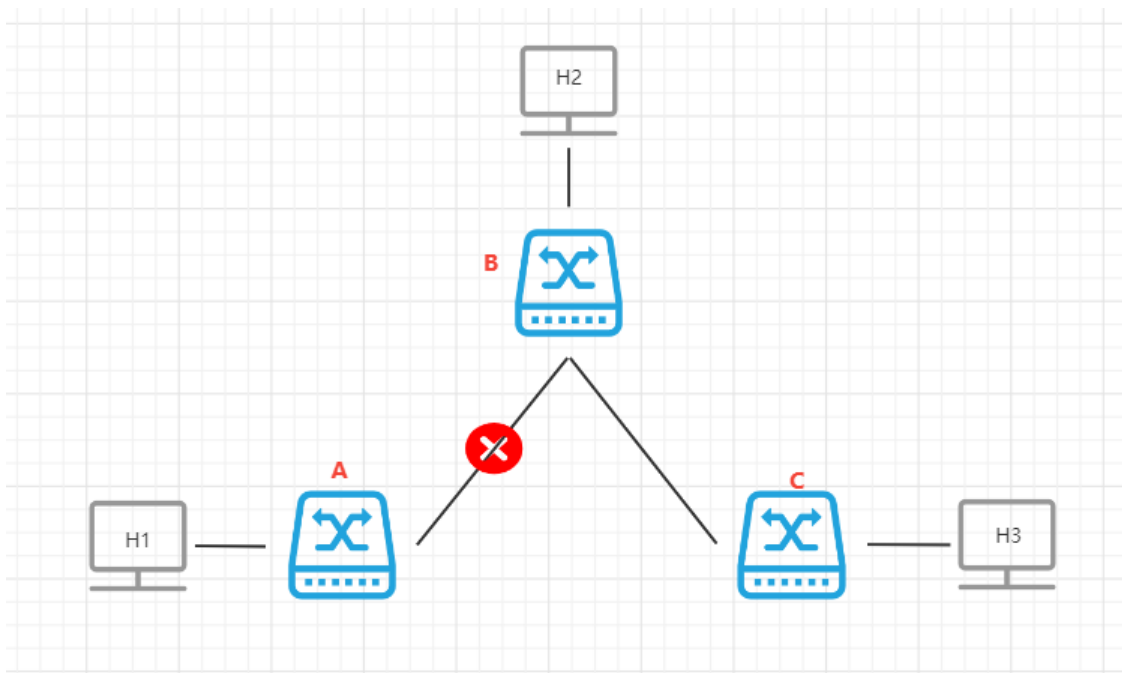
一、如何提高以太网可靠性

我们来看看问题的由来。

例如下图所示的以太网由三台交换机互连而成，每个交换机上都连接着一些主机：



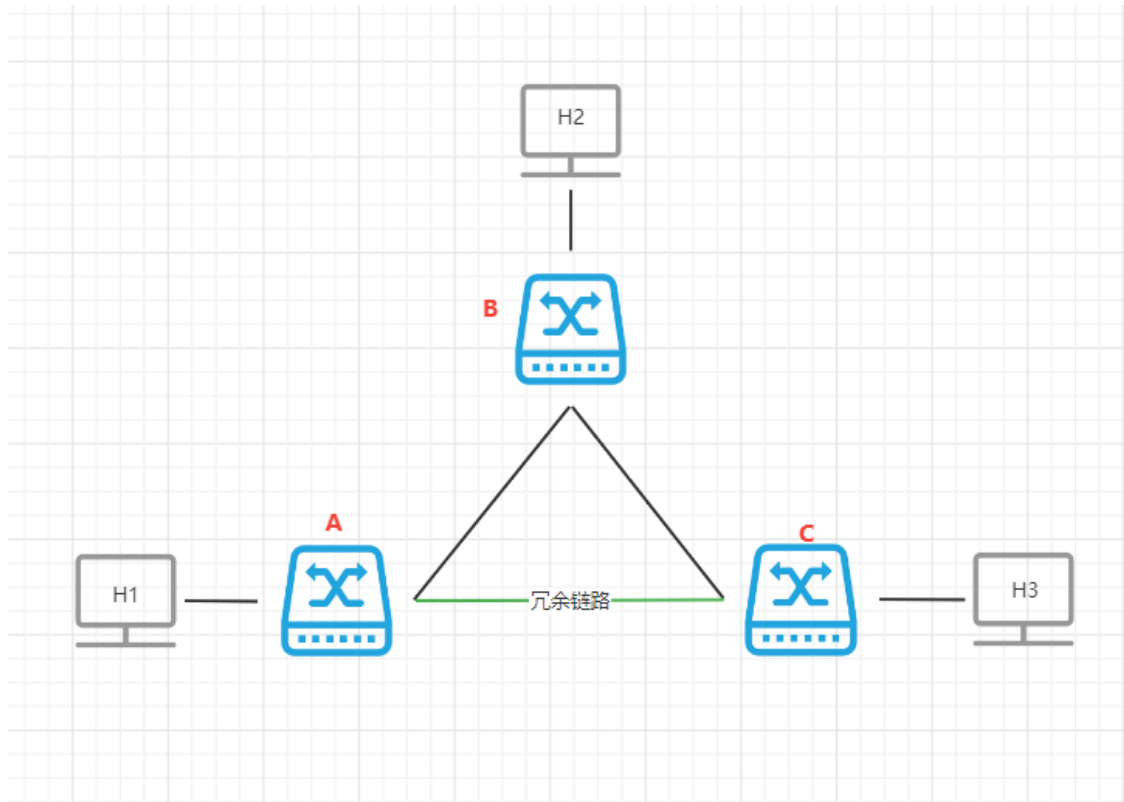
此时若其中一条链路发生故障，比如A和B之间的链路发生故障：



此时，交换机B上的所有主机跟交换机A和C上所连接的主机们都无法通信。更加极端时，当AB和AC之间的链路都发生故障，那么这个结构将被拆分为三个独立的交换局域网，之间都互相不能通信了。

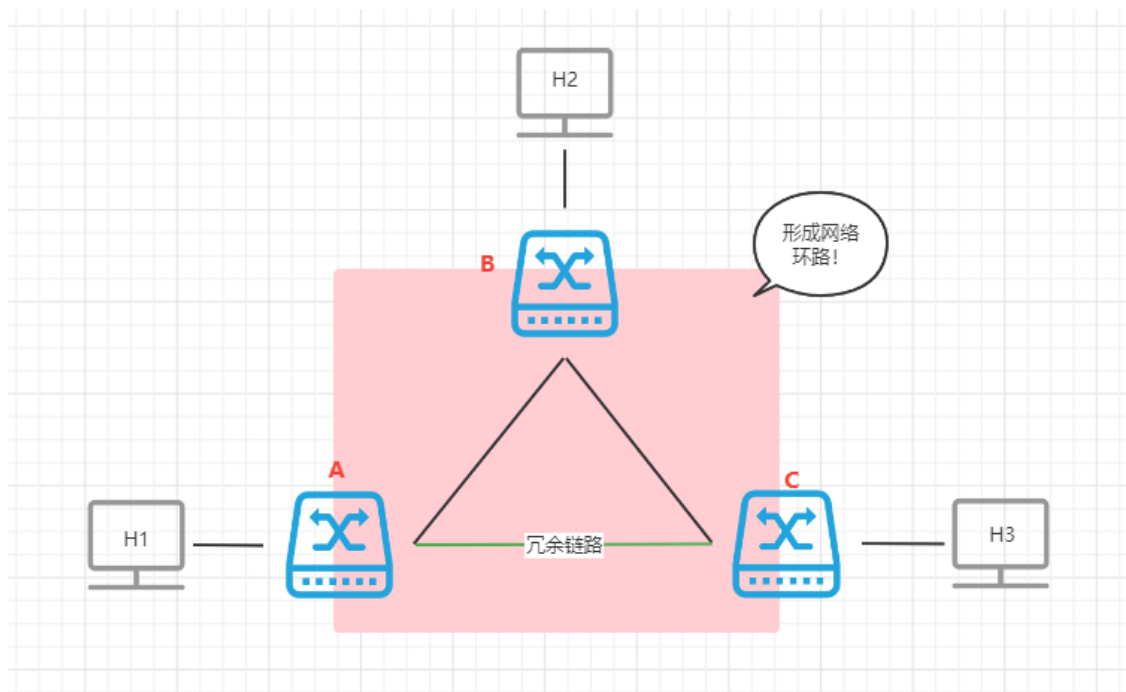
此时我们如何才能提高这种以太网交换网络的可靠性呢？相信大家很容易想到一个思路：通过添加冗余链路的方法来提高以太网的可靠性。

下面我们来加一根冗余链路：



此时即使A和B之间的链路发生了故障，三台交换机之间还是依然有链路让他们互通，一定程度上提高了可靠性。

没错，引入冗余链路即可有效提高以太网的可靠性，但是它是一把双刃剑，会带来负面效应：形成网络环路。



二、网络环路

你可能会问，网络环路就环路了呗，会有啥影响呢？不，这影响老大了，下面我们来逐一说明。

第一个问题：网络环路会造成广播风暴。

试想，假设此时主机H1发送一个广播帧，会发生什么。

首先交换机A收到广播帧后，需要将其转发，转发给交换机B和C；

交换机C收到来自A的广播帧后，需要将其转发，转发给交换机A和B（B在想：你有病？我不刚收到吗？）；

交换机B收到来自A的广播帧后，需要将其转发，转发给交换机A和C（C在想：你有病？我不刚收到吗？）；

这样下去，A又将收到同样的广播帧（A在想：你有病？我不刚收到吗？）, 只不过是来自交换机B和C的；B也将再次收到同样的广播帧，来自交换机A和C的，交换机C同理。

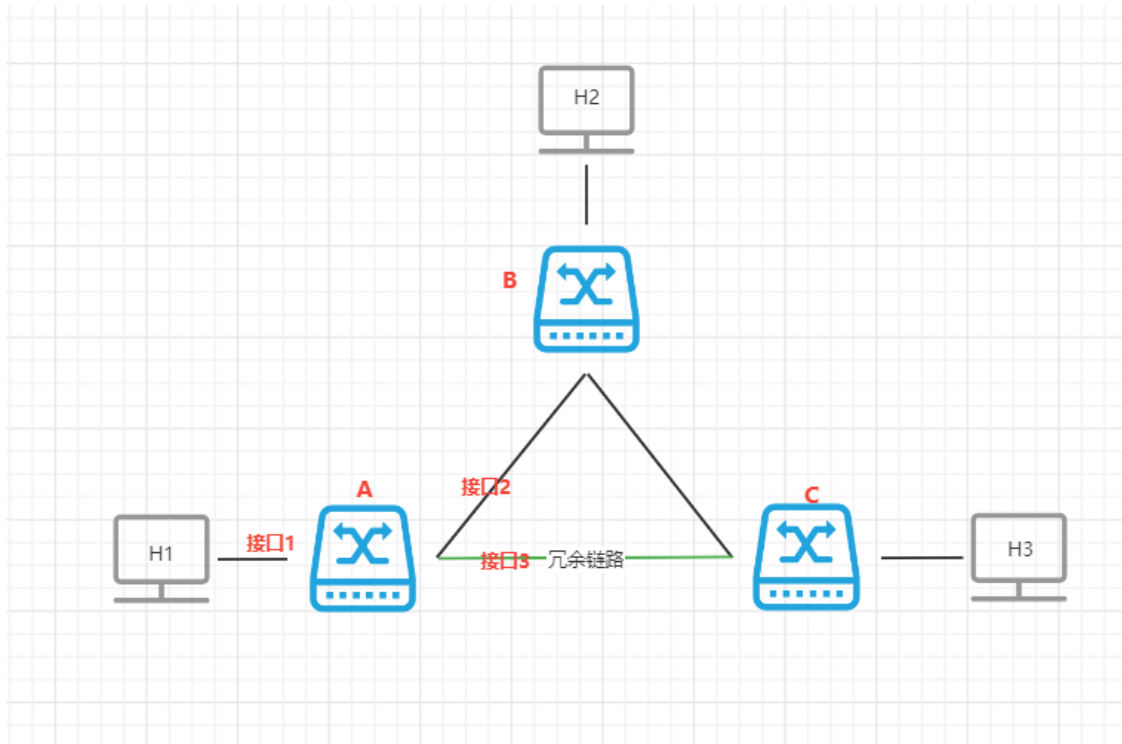
这样，交换机B将不断收到交换机A和C发来的广播帧，并且将其转发；交换机C又将不断收到交换机A和B的广播帧并将其转发；交换机A不断收到交换机B和C发来的广播帧并将其转发，如此反复陷入死循环...

广播风暴是指局域网中的设备大量发送广播报文,而交换机无法对正常数据进行转发的现象。当网络中出现拥塞现象时(如交换机CPU负载偏高),需要使用抓包工具分析网络中广播包的比例是否过高,以确定是否发生了广播风暴。

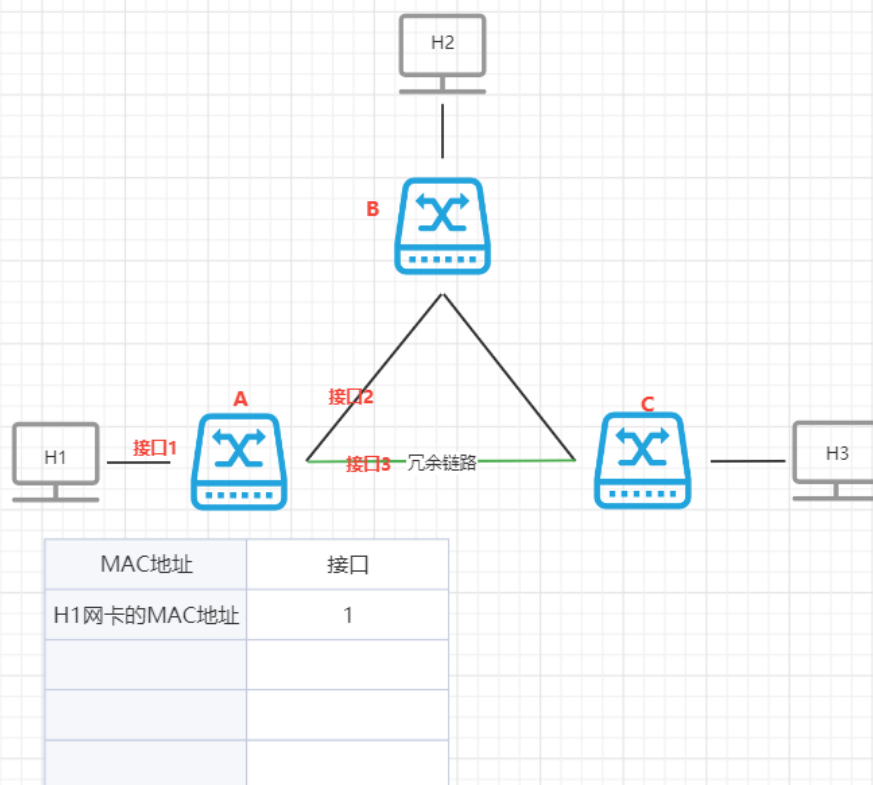
然而要注意,交换机环路并不是引起广播风暴的唯一原因,ARP病毒、蠕虫病毒等都能引起广播风暴,在发生广播风暴时,不能简单地认为网络中出现了交换机环路。

第二个问题:造成交换机的帧交换表震荡(漂移)。

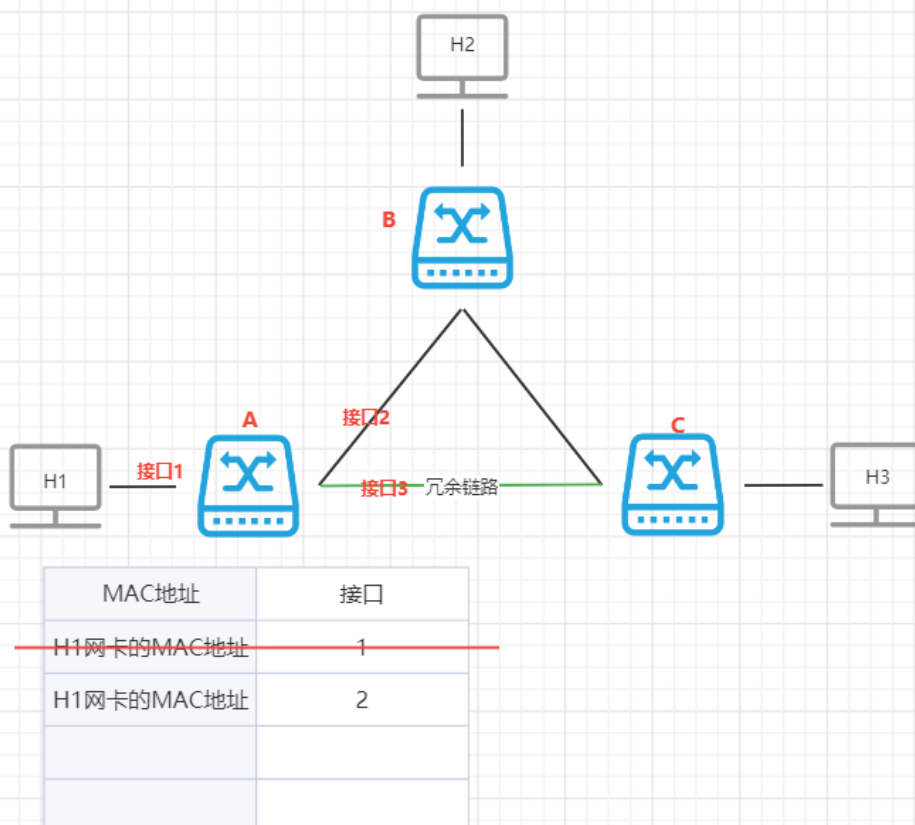
我们接着上面的三台交换机组成的以太网网络为例,假设主机H1连接交换机A的接口1,交换机B连接的是交换机A的接口2,交换机C连接的是交换机A的接口3。



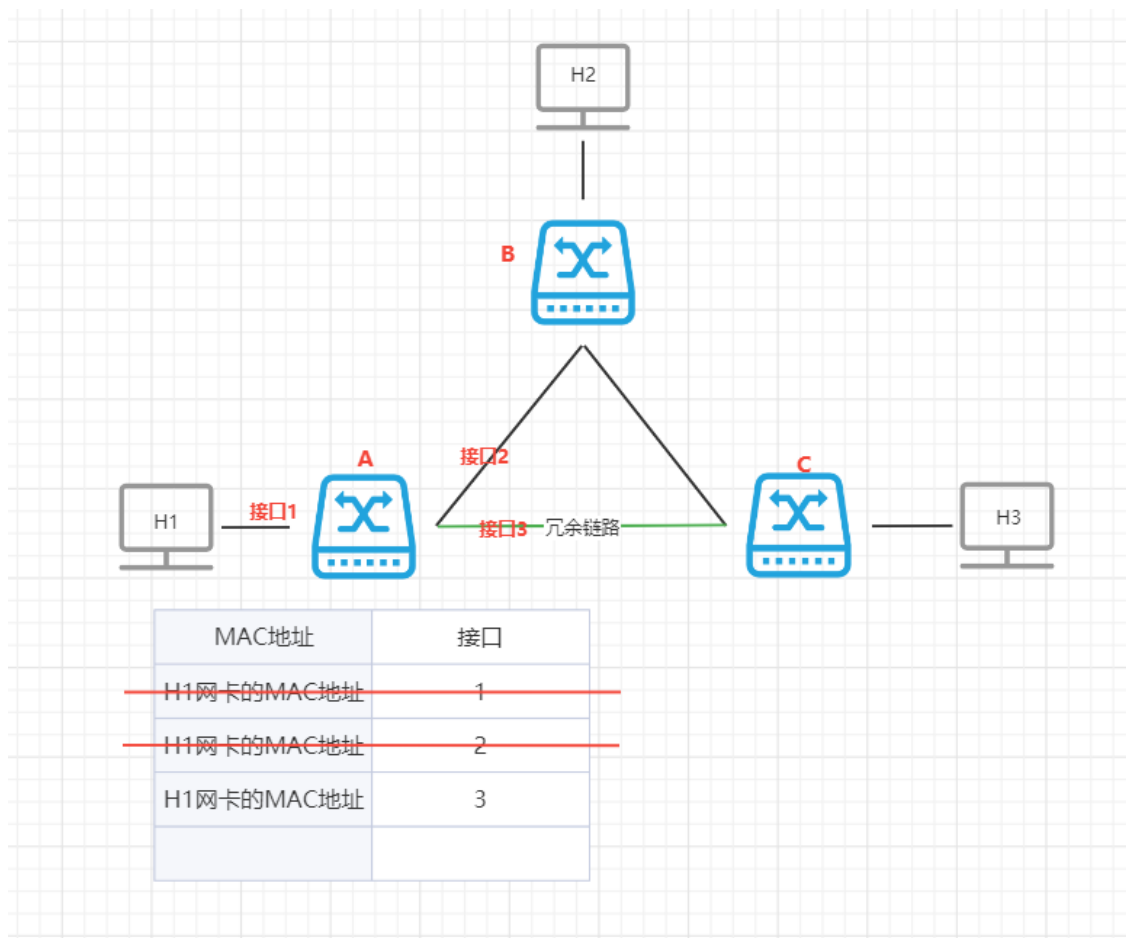
当主机H1发送一个广播帧后,交换机A将会在CAM表中记录如下:



此时记录还是正确的，但是由于发生了广播风暴，不久之后将会收到交换机B转发来的广播帧，此时为接口2，此时在CAM表进行更新：



此时记录实际上已经被错误更新了，不过此时又收到交换机C转发来的广播帧，接口为3，又要对CAM表进行更新：



记录又被错误地更新了一次，后续，将还会重复收到交换机B和交换机C转发来的广播帧，CAM表记录将会在2和3之间来回横跳，即反复震荡或者说反复漂移。

三、生成树协议STP、网络环路

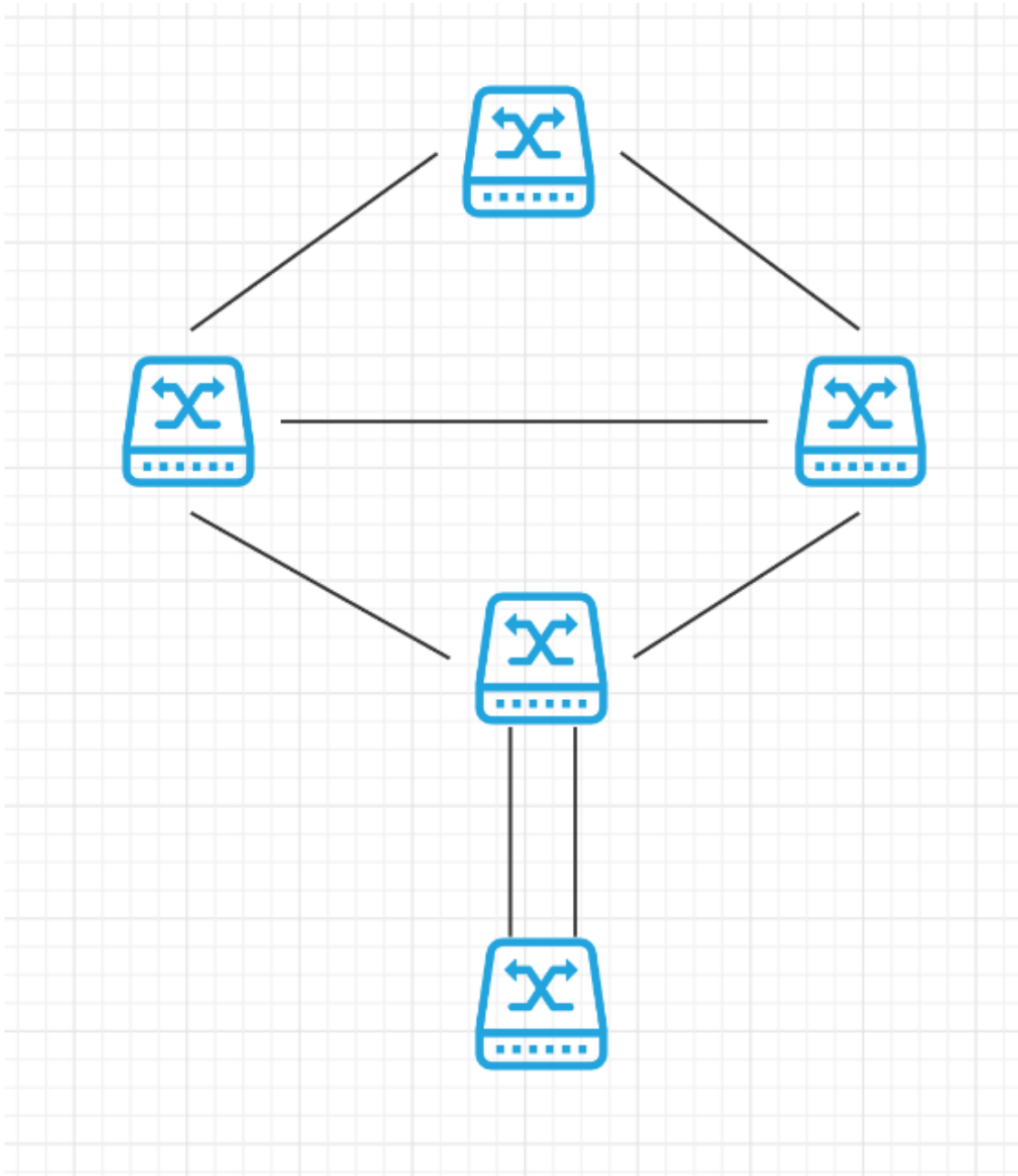
如何既可以让网络环路起到他的提高以太网网络可靠性的目的，又能避免由于链路冗余导致网络环路继而引发广播风暴和帧交换表震荡等问题呢？

以太网交换机使用生成树协议STP (Spanning Tree Protocol) ，可以在增加冗余链路来提高网络可靠性的同时又避免网络环路带来的各种问题。

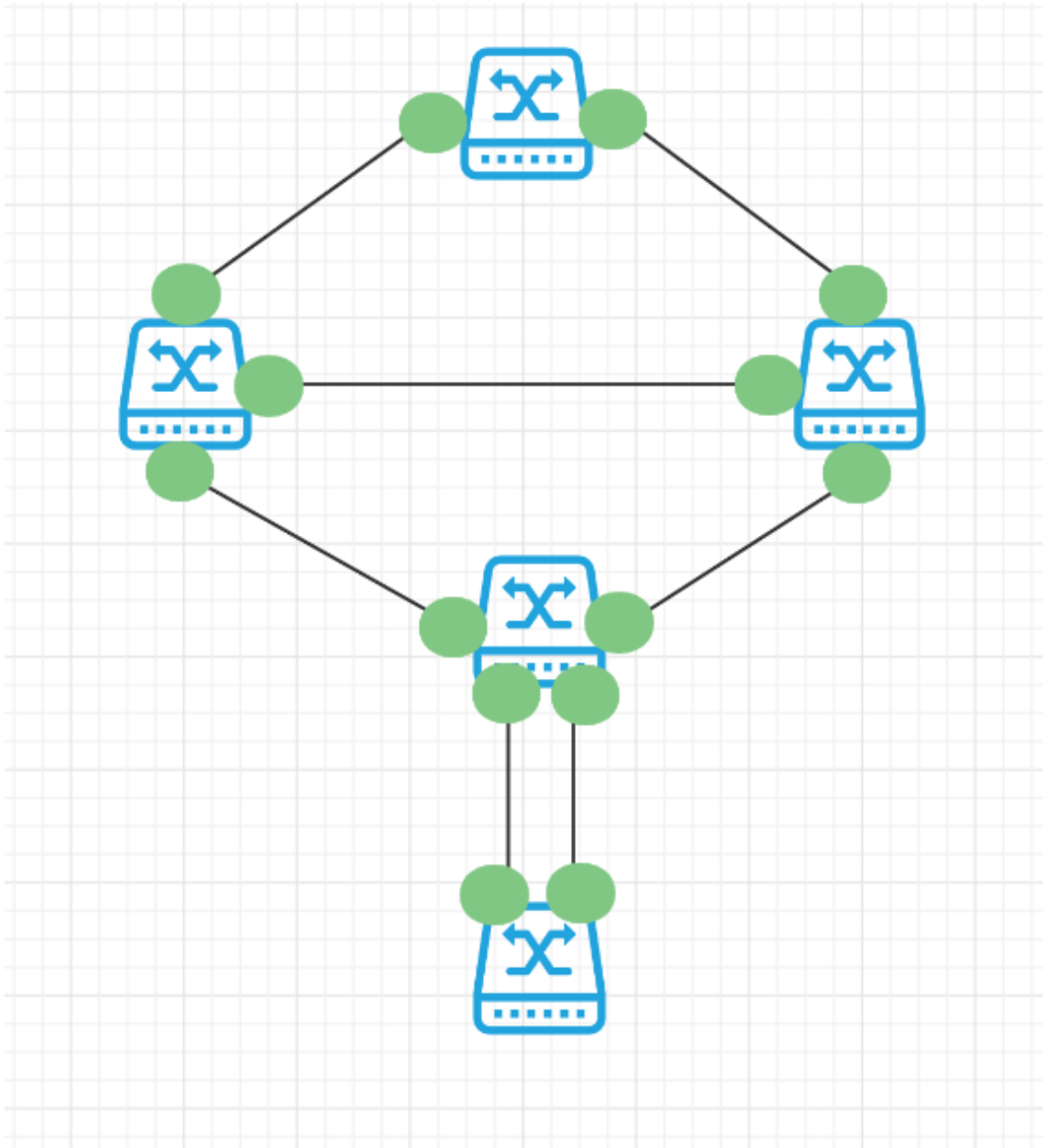
无论交换机之间采用怎样的物理连接，交换机都能够自动计算并构建一个逻辑上没有环路的网络，其逻辑拓扑结构必须是树型的（无逻辑环路）。

实际上各交换机之间按照生成树协议规定的生成树算法，交互一些参数后，就可以判断出自己应该阻塞自己的哪些接口。

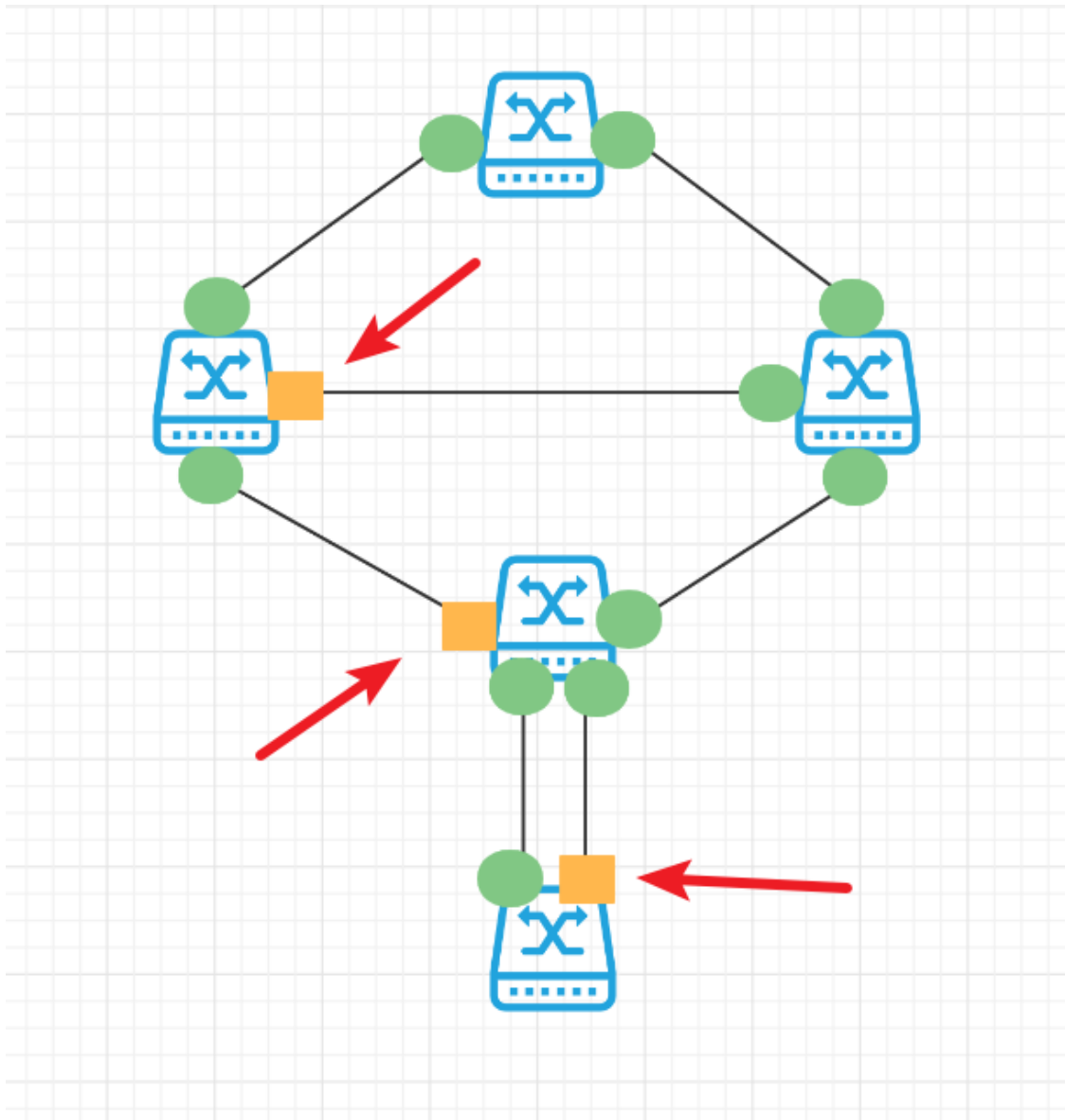
下面我们结合具体的例子来说明，如图所示是由五台交换机构建的以太网，并且环路还不止一个。



在没有结合生成树算法前，每个接口都是互通的，我们用绿色的小圆圈来表示目前的交换机链路互通情况：



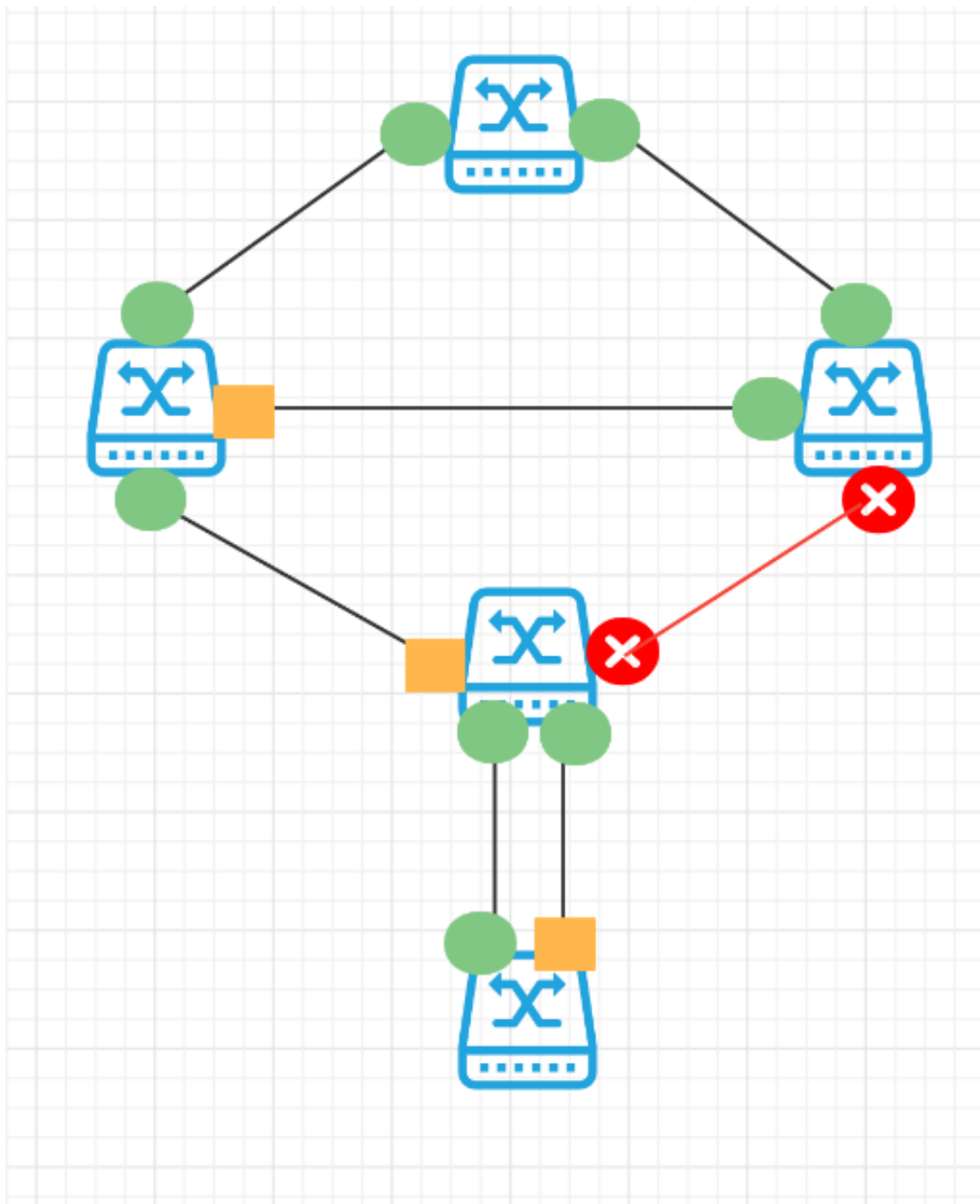
此时，交换机结合生成树协议，交互一些参数后，判断出来自己应该阻塞哪些接口，例如下图所示：



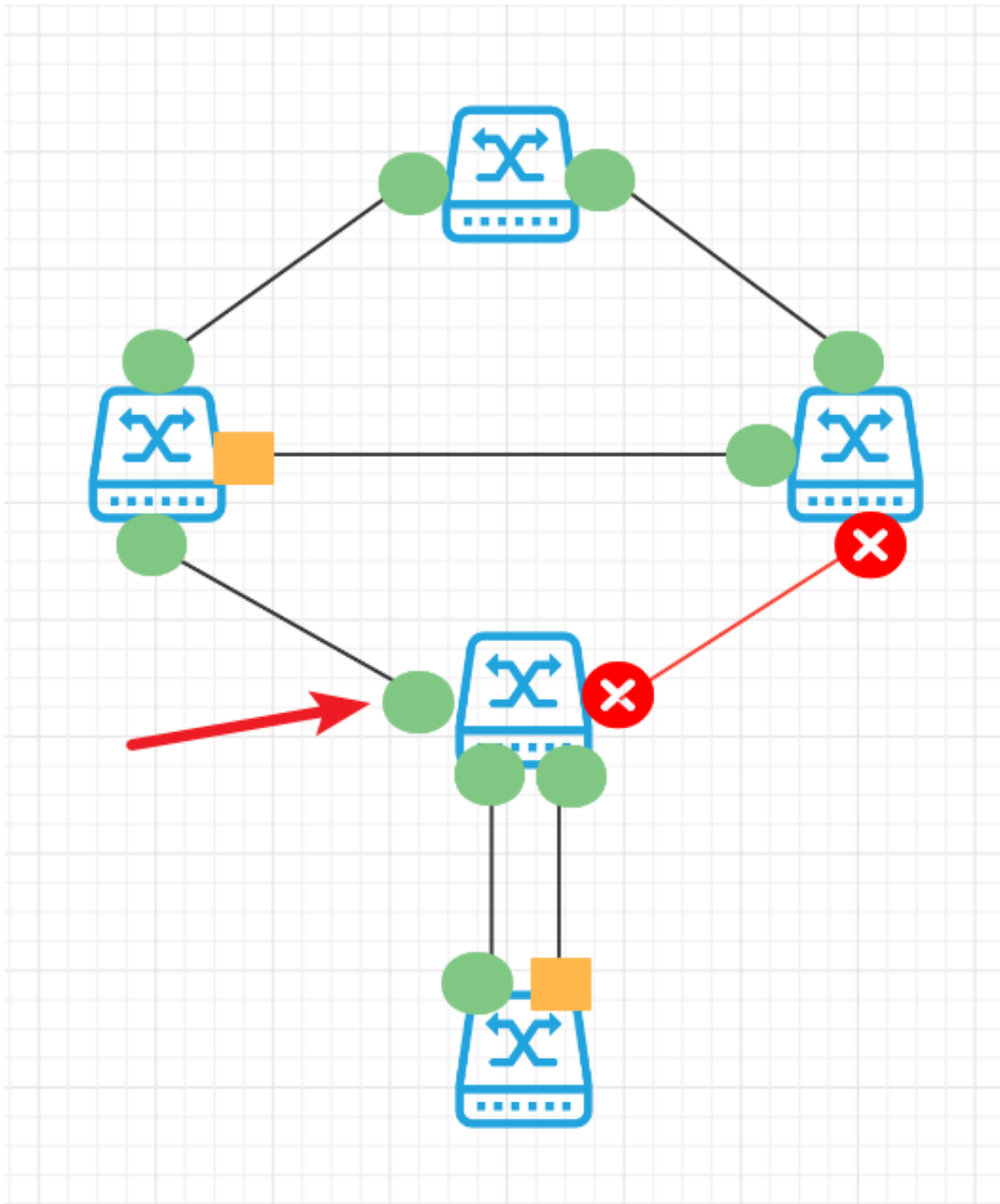
这样就形成了一个逻辑上没有环路的网络，并且也确保了最终生成的树型逻辑拓扑连通整个网络。

当首次连接交换机或网络物理拓扑发生变化时（认为改变或故障），交换机都将进行生成树的重新计算。

假设某条链路发生故障导致网络不通：



相关交换机检测到该故障后，重新计算生成树，决定将自己之前阻塞的接口恢复为正常状态，形成一个新的逻辑上没有环路，但是又连通所有交换机的网络。



关于生成树算法的细节不在这里进行学习和说明，不是本系列关注的重点。下一篇文章将学习虚拟局域网VLAN技术。