**多生成树MSTP的理解**

　　多生成树协议MSTP（Multiple Spanning Tree Protocol）是IEEE 802.1s中定义的一种新型生成树协议,它引入了“实例”（Instance）的概念。简单的说，STP/RSTP是基于端口的，PVST＋是基于VLAN的，而MSTP就是基于实例的。所谓实例就是多个VLAN的一个集合，通过多个VLAN捆绑到一个实例中去的方法可以节省通信开销和资源占用率。MSTP各个实例拓扑的计算是独立的，在这些实例上就可以实现负载均衡。在使用的时候可以把多个相同拓扑结构的VLAN映射到一个实例里，这些VLAN在端口上转发状态将取决于对应实例在MSTP里的状态。MSTP的实例0具有特殊的作用，称为CIST，即公共与内部生成树，其他的实例称为MSTI，即多生成树实例。

MSTP协议引入了域的概念，域由域名、修订级别、VLAN与实例的映射关系组成，只有三者都一样的互联的交换机才认为在同一个域内。缺省时，域名就是交换机的第一个MAC地址，修订级别等于0，所有的VLAN都映射到实例0上。在同一个域内的交换机将互相传播和接收不同生成树实例的配置消息，保证所有生成树实例的计算在全域内进行；而不同域的交换机仅仅互相传播和接收CIST生成树的配置消息，MSTP协议利用CIST保证全网络拓扑结构的无环路存在，也是利用CIST保持了同STP/RSTP的向上兼容，因此从外部来看，一个MSTP域就相当于一个交换机，对不同的域、STP、RSTP交换机是透明的。

　　MSTP相对于之前的各种生成树协议而言，优势非常明显。MSTP具有VLAN认知能力，可以实现负载均衡，可以实现类似RSTP的端口状态快速切换，可以捆绑多个VLAN到一个实例中以降低资源占用率，并且可以很好地向下兼容STP/RSTP协议。

　　MSTP中的几个关键技术点：

　　l 端口状态

　　STP协议通过在交换机之间传递特殊的消息（这个消息称为BPDU，桥协议数据单元，又成为配置消息Configuration Message）,并进行分布式的计算，来决定一个有环路的网络中，哪台交换机的哪个端口应该被阻塞（Discarding），用这种方法来剪切掉环路。称这种被阻塞的端口为：其处在DISCARDING状态，处于DISCARDING状态的端口不会转发任何报文，只会接收特定的BPDU报文，对于其他的报文一律丢弃。需要注意的是，端口的状态是一个瞬时值，并非一成不变的.

　　除了处于DISCARDING状态的被阻塞的端口以外，还有一种正常转发数据报文的端口，其状态称为FORWARDING，与不启动MSTP协议的正常端口没有什么分别。

　　除了FORWARDING、DISCARDING、LEARNING状态以外，还有一种端口状态DISABLE，是指物理上LINK DOWN的端口和没有启动STP协议的端口。

　　l 根桥和上游桥

　　根桥就是那棵生成树的总根，整个网络中有且只有一个根桥。它是整个网络的逻辑中心，但不一定是物理中心。当MSTP协议开始运行时，就开始了选举根桥的过程，各交换机之间通过比较各自的桥ID来选举根桥，最后整个网络中桥ID最小的交换机成为根桥。桥ID（类似于OSPF中的Router ID）它由交换机的优先级+MAC地址构成，比较的时候先比较优先级，如果优先级相同则比较MAC，比较的原则都是值较小者优。

我们知道，树是一个由上至下分层的结构，比如WINDOWS的文件系统，而在MSTP中，总根就是最高的那一层“我的电脑”，各个盘符C、D、E相当于是它的子目录，而在MSTP中就称之为下游桥，总根就是下游桥的“上游桥”。可见一个上游桥可能会有多个下游桥，而一个下游桥，有且只有一个上游桥。那么这里是怎么判断一个交换机的上游桥或下游桥呢？在MSTP中的判断方法就是COST值（类似路由协议中的METRIX），也叫做开销，COST值从根桥时为0，在每经过一个接收端口就会累加该端口的COST。在非根桥交换机中距离总根桥开销最小的那个端口被称作Root端口，就是根端口。一台交换机在一个实例上有且只有一个根端口，根端口对端相连接的交换机就是这台交换机的上游桥。

　　l 端口角色

　　在上一节中所描述的“根端口”其实就是说的端口角色。MSTP中共有四种端口角色，分别为：Designed、Root、Alternate、Backup；简称为DESI、ROOT、ALTE、BACK。

　　端口角色和端口状态之间存在一定的联系，一般情况下，DESI和ROOT端口的状态都为FORWARDING（LEARNING）；而ALTE和BACK端口的状态永远为DISCARDING。

　　这样就很好理解了，根端口的角色是怎样选举产生的，ALTE端口和BACK端口其实就是我们所说的被阻塞的端口，而与ALTE、ROOT、BACK端口相连接的对端就是DESI端口（对端只有一个端口的情况下）。在网段上抑制其他端口（无论是自己的还是其他网桥的）发送BPDU的端口，就是该网段的指定端口。在收敛后，只有指定端口和根端口可以处于转发状态。这个定义比较难懂，简单说来，MSTP启动以后，每个端口都会发送自己的配置消息，这个消息是端口当前认为的最优消息，此时就会发生比较。如果某端口发送的消息优于对端，那么此端口就是DESI端口，对端则为其他三种端口角色之一。

　　MSTP协议从开始运行起，一边发送消息通过比较产生根桥，同时根端口和DESI端口也被选举出来，剩下的端口就是ALTE和BACK端口。

　　l 域与多实例

域就拥有相同域信息的一些相连通的交换机所组成的区域，可以理解为一个自治系统。MSTP中的实例与VLAN采取“映射”的概念，一个实例可以对应多个VLAN，而一个VLAN只能对应一个实例。几个交换机处于同一个域内，就是他们的域信息完全一致：包括域名，格式前缀（一般都为0不会有什么变化），修订级别，VLAN和实例之间的映射关系。

　　在MSTP的生成树算法中，一个域被当作一台交换机来处理，这样计算生成的树叫做CST（公共生成树）。每个域内同样运行MSTP协议，对每一个实例都计算出一棵生成树，叫做MSTI（多生成树实例），MSTI是域内才有的概念。此外域内的实例0是一个特殊的实例，实例0的生成树和CST一起组成了一个包含全网内所有交换机的树，叫做CIST（公共内部生成树）。域和域之间的数据流量通过CIST来实现，域内则根据MSTI来实现，从而实现了RSTP与VLAN的完美结合。

　　虽然多了域与实例，其实并不复杂，在全局来看，一个域就是一台交换机，和刚才介绍的算法没有什么本质区别。而在域内，就相当于是一个运行RSTP的小型网络，只不过根桥被域根桥所取代。域内各MSTI之间各自独立计算自己的生成树，互相之间互不干扰。

　　域也将路径开销分为了内部路径开销和外部路径开销. 内部路径开销也称IRPC,指本交换机到所在域的域根桥的开销；外部路径开销也称ERPC，是指从域根到总根的路径开销。在MSTP中选举指定端口和根端口时，比较的先后顺序依次为：根桥ID，外部路径开销，域根ID，内部路径开销，指定桥ID，指定端口ID，接收端口ID 。