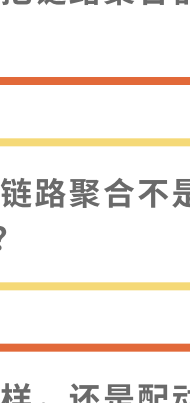


交换机网络可靠性小贴士

H3C交换机基础维护篇

兄弟，周末一起打游戏呀~

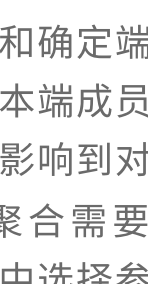


不了老哥，这周末我要给客户做变更，得把链路聚合都改成动态的。

为啥呀？链路聚合不是动态静态都可以吗？

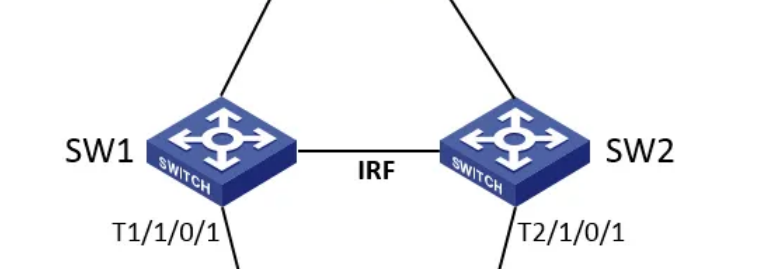
这可不一样，还是配动态好，前两天这个局点出了大问题，就是因为静态链路聚合切业务没切成功。

还有这事儿？快给我讲讲具体咋回事儿！



小贴士No.1 配置动态链路聚合

链路聚合想必大家都很熟悉了，链路聚合将多条物理链路捆绑在一起形成一条逻辑链路，实现增加链路带宽的目的，同时这些链路相互备份，可以有效地提高链路的可靠性。



新华三交换机的链路聚合分为动态链路聚合和静态链路聚合：

静态聚合模式：一旦配置好后，端口的选中/非选中状态就不会受网络环境的影响，比较稳定。

动态聚合模式：通过LACP协议实现，能够根据对端和本端的信息调整端口选中/非选中状态，比较灵活。

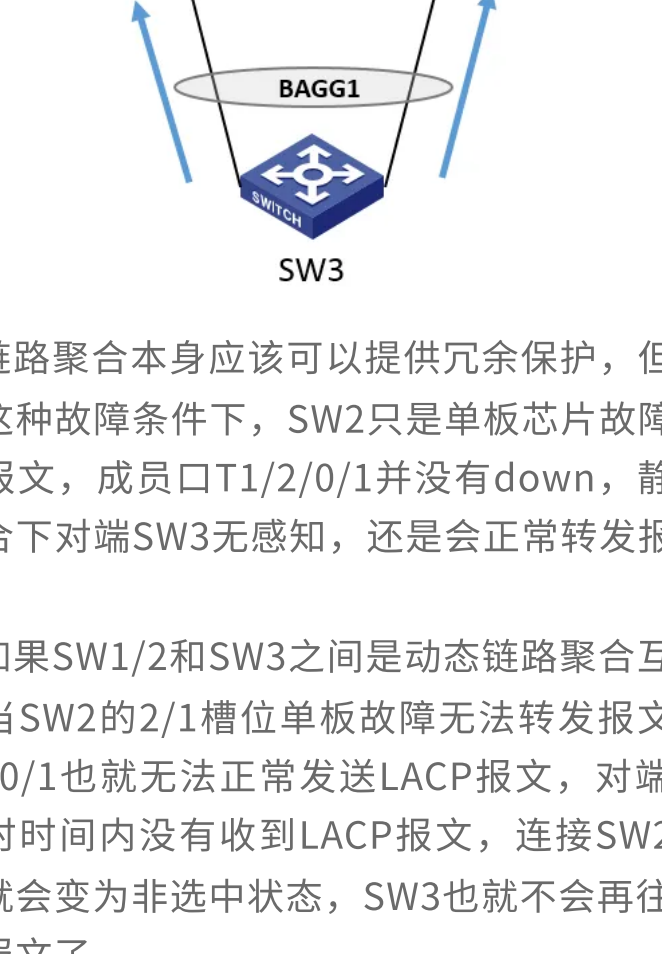
动态链路聚合和静态链路聚合最大的差别就在于，在选择参考端口和确定端口选中状态时，静态链路聚合只需要比较本端成员端口，本端成员端口属性类配置变化不会影响到对端的选中状态，比较稳定；而动态链路聚合需要和对端交互LACP协议，在两端成员端口中选择参考端口，端口的选中状态同时受本端和对端影响。

PS：关于链路聚合的具体介绍，大家可以参考新华三官网的配置手册，这里我们就不赘述啦~

交换机缺省是静态链路聚合，动态链路聚合需要多配置一条命令(**link-aggregation mode dynamic**)，在网络稳定情况下，静态和动态两种模式没有什么差别，甚至静态链路聚合下设备还不需占用CPU处理LACP协议报文；但在有些故障环境下，静态链路聚合就有问题了。我们来看一个故障案例。

· 组网说明 ·

现场使用一组堆叠框式设备作为核心，如下图中SW1和SW2，通过跨框静态链路聚合下联一台汇聚设备。

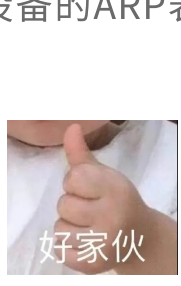


· 故障现象 ·

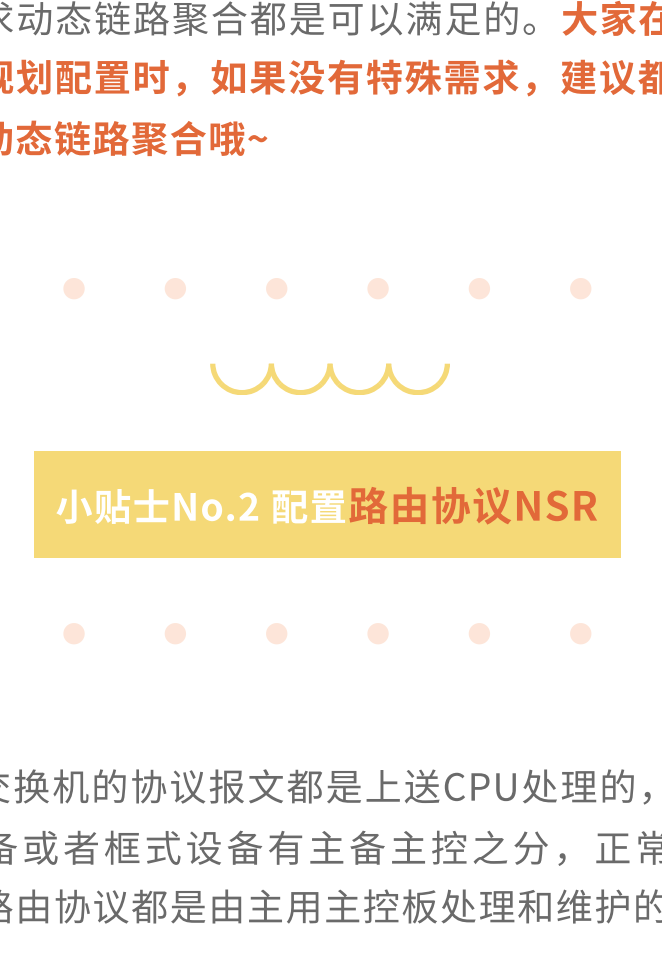
网管监控平台上收到大量核心交换机2/1槽位单板芯片异常告警，与此同时下行业务也开始报障，现场运维工程师将聚合组1中的T2/1/0/1成员口手动shutdown后，业务恢复正常。

· 故障分析 ·

根据核心交换机上2/1槽位的告警日志，可以确认2/1槽位单板芯片出现了硬件故障，影响了所有通过2/1槽位转发的业务。但是现场核心下行到汇聚SW3是通过聚合组互联的，且是跨框聚合，两个成员口分别在SW1的1/1槽位单板和SW2的2/1槽位单板上，即使2/1槽位单板故障了，流量也可以走1/1槽位单板转发，不应该影响业务呀。

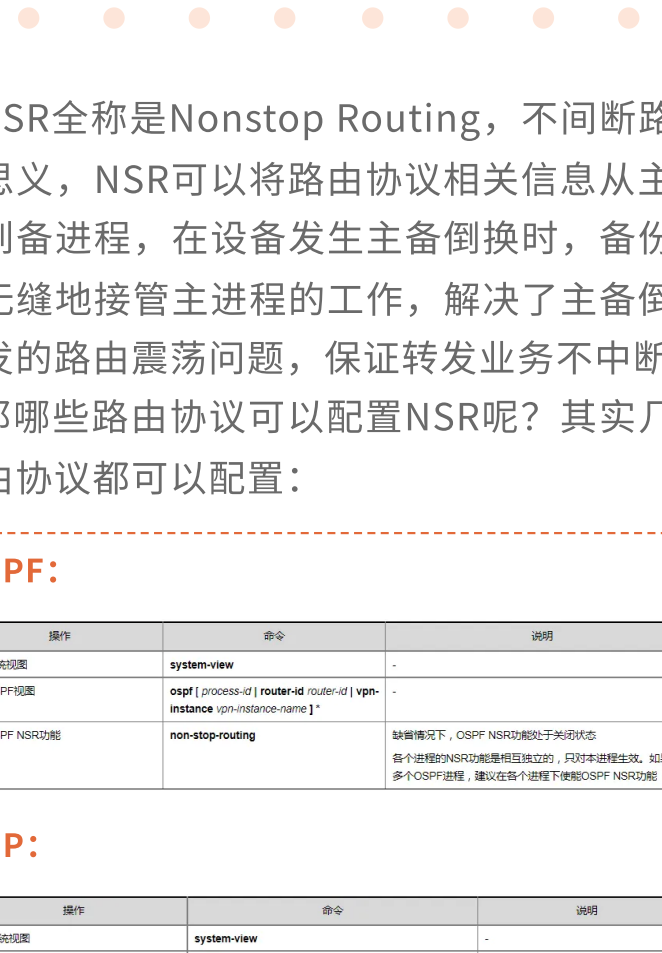


经过进一步分析，发现在故障时间点，下行SW3上聚合组1的成员口还都是选中状态，所以还是会将流量hash到SW1和SW2上，hash到SW2的报文由于单板故障无法转发，就全部被丢弃了，导致影响业务：



链路聚合本身应该可以提供冗余保护，但是在现场这种故障条件下，SW2只是单板芯片故障无法转发报文，成员口T1/2/0/1并没有down，静态链路聚合下对端SW3无感知，还是会正常转发报文到SW2。

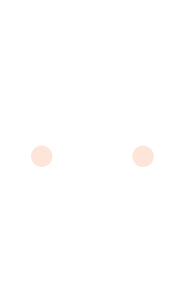
如果SW1/2和SW3之间是动态链路聚合互联，那么当SW2的2/1槽位单板故障无法转发报文时，T1/2/0/1也就无法正常发送LACP报文，对端SW3在超时时间内没有收到LACP报文，连接SW2的成员口就会变为非选中状态，SW3也就不会再往SW2发送报文了。



除了以上案例，在很多其他故障场景下，静态链路聚合也无法及时切换隔离故障：

比如两台设备间有传输设备的情况，其中一台的端口已经down了，但另一台的端口还是up的，静态链路聚合依旧会保持选中状态；

比如一端为堆叠跨框聚合，单框设备故障导致堆叠分裂且没有及时合并的情况下，对端还是会保持选中，不仅会将报文转发给故障设备，还会学习故障设备的ARP表项，造成严重的业务影响。



静态链路聚合不关注对端的特性让它更稳定，但也无法及时感知对端故障，只要成员口还是UP的，就会保持选中状态，将业务流量持续发送对端；而动态链路聚合实时交互LACP协议报文，可以及时感知到对端的异常，我们平时遇到的绝大部分单链路故障，都可以通过配置动态链路聚合实现冗余备份。

虽然动态链路聚合也有其局限性，无法满足毫秒级的切换要求，但大部分普通组网的链路冗余备份需求动态链路聚合都是可以满足的。**大家在开局组网规划配置时，如果没有特殊需求，建议都尽量使用动态链路聚合哦~**

小贴士No.2 配置路由协议NSR

交换机的协议报文都是上送CPU处理的，而堆叠设备或者框式设备有主备主控之分，正常运行时，路由协议都是由主用主控板处理和维持的：

Slot Type	State	Subslot	Soft Ver	Patch Ver
1 56800-4C	Master	2	2432P03	None
2 56800-4C	Standby	2	2432P03	None

Slot Type	State	Subslot	Soft Ver	Patch Ver
1/0	Normal	0	S10500-7595P02	None
1/1	NONE	Absent	0	NONE
1/2	NONE	Absent	0	NONE
1/3	NONE	Absent	0	NONE
1/4	NONE	Absent	0	NONE
1/5	NONE	Absent	0	NONE
1/6	LSUM1MPU06B0	Standby	0	S10500-7595P02
1/7	LSUM1MPU06B0	Master	0	S10500-7595P02
1/8	LSUM1FAB06C0	Normal	0	S10500-7595P02
1/9	LSUM1FAB06C0	Normal	0	S10500-7595P02

当主用主控板重启或者故障时，设备会发生主备倒换，由备用主控接管路由协议，过程中涉及邻居关系的重新建立和路由的重新收敛，可能会导致路由协议震荡，进而影响业务。

那有没有什么办法减少或者避免主备倒换时的路由震荡呢？有的，答案就是配置路由协议NSR~

NSR全称是Nonstop Routing，不间断路由。顾名思义，NSR可以将路由协议相关信息从主进程备份到备进程，在设备发生主备倒换时，备份进程能够无缝地接管主进程的工作，解决了主备倒换期间引发的路由震荡问题，保证转发业务不中断。

那哪些路由协议可以配置NSR呢？其实几乎所有路由协议都可以配置：

操作	命令	说明
进入系统视图	system-view	-
进入OSPF视图	ospf [process-id] [router-id router-id vpn-instance { vpn-instance-name }]	-
使能OSPF NSR功能	non-stop-routing	缺省情况下，OSPF NSR功能处于关闭状态。各个进程的NSR功能相互独立的，只对本进程生效。如果存在多个OSPF进程，建议在每个进程下使能OSPF NSR功能。

操作	命令	说明
进入系统视图	system-view	-
进入BGP视图	bgp [as-number instance { instance-name }]	-
使能BGP NSR功能	non-stop-routing	缺省情况下，BGP NSR功能处于关闭状态。各个进程的NSR功能相互独立的，只对本进程生效。如果存在多个BGP进程，建议在每个进程下使能BGP NSR功能。

操作	命令	说明
进入系统视图	system-view	-
进入ISIS视图	isis [process-id] [vpn-instance { vpn-instance-name }]	-
使能ISIS NSR功能	non-stop-routing	缺省情况下，ISIS NSR功能处于关闭状态。各个进程的NSR功能相互独立的，只对本进程生效。如果存在多个ISIS进程，建议在每个进程下使能ISIS NSR功能。

· NSR注意事项 ·

1、NSR特性与GR特性互斥，不能同时配置，如果同时配置会有如下报错：

配置了NSR后，再配置GR就会有报错： [H3C-ospf-1] graceful-restart OSPF can't configure GR when NSR configuration is present.
配置GR后再配置NSR后也会报错： [H3C-ospf-1] non-stop-routing OSPF can't configure NSR when GR configuration is present.

2、在堆叠升级/更换、更换主控板等涉及到主备倒换的操作之前，务必先配置NSR，减少变更过程中的业务影响。

3、交换机基本所有设备都支持NSR，建议联系400或者代表处确认并升级到推荐版本，NSR优化更平滑~

断网让人胆战心惊，多少次处理到天明，蓦然回首，原来只需一条命令。

动态链路聚合和NSR的配置都很简单，却可以避免大问题，快检查一下自己的交换机上有没有配置吧~

END

扫码关注我們哦