



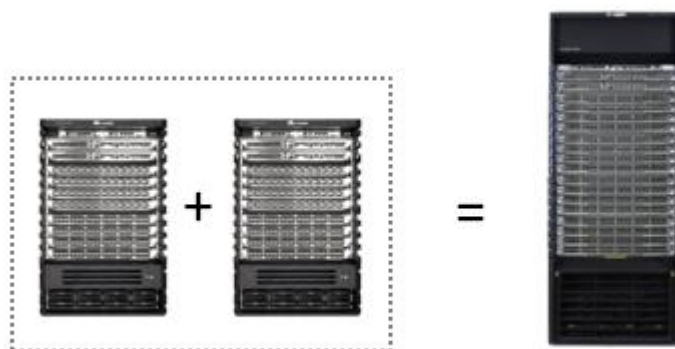
浅谈堆叠与M-LAG

信锐网技术有限公司

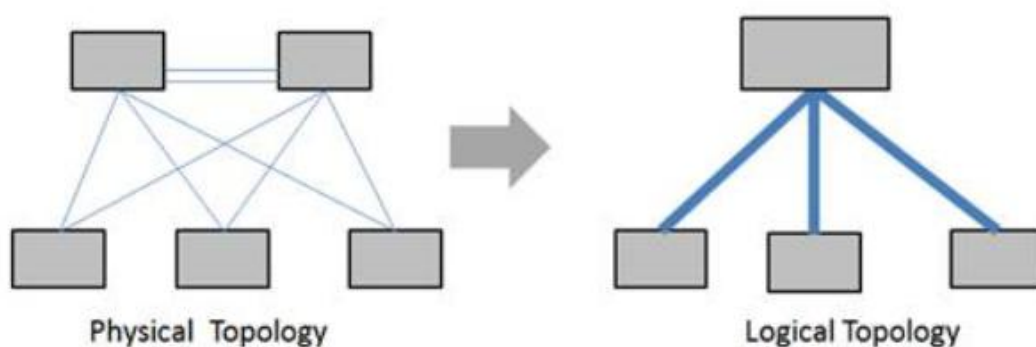
2018年12月11日

一、堆叠

堆叠技术，本质上就是合并，管理平面、控制平面、转发平面的全面合并。堆叠系统的主控板，管理两台物理设备的所有线卡和网板，变成一个逻辑的大交换机



当然堆叠目的不仅仅是为了变大，从网络角度看一下逻辑拓扑，一下变得“高富帅”！



”高富帅”的主要表现：

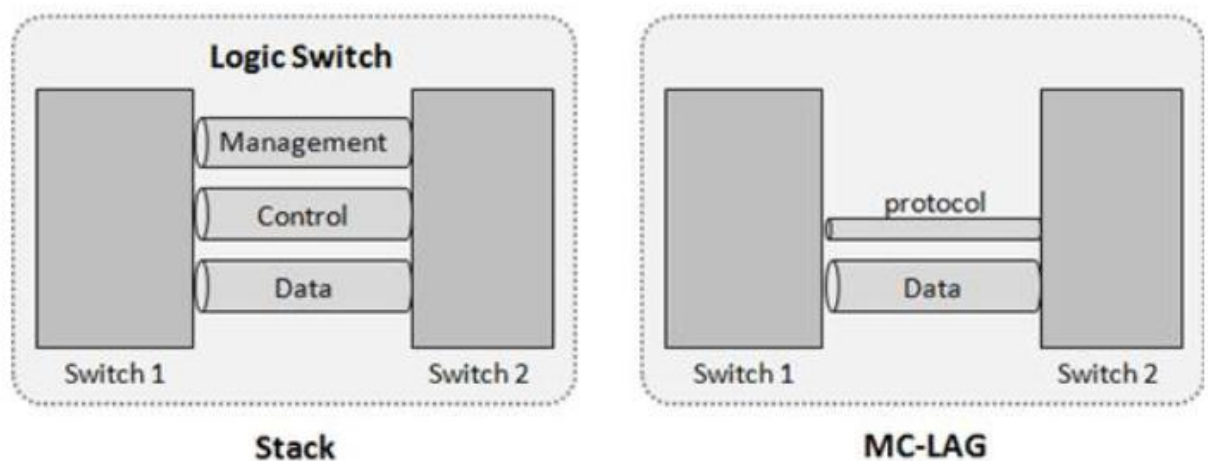
- 几乎两倍交换能力的超级节点；
- 二三层转发流量完全负载分担，充分利用所有链路；
- 逻辑单节点，业务支持全面，网络方案设计简单

- 通过部署跨框link-agg，支持物理节点的故障保护；
- 网元二合一，有利于网络管理和维护。
- 相对传统STP，可以组建更大的二层网络；
- link-agg的收敛性能，网络故障收敛快。

二、Link-agg虚拟化 (M-LAG)

MC-LAG (Multi-Chassis Link Aggregation Group)，CISCO称之为vPC

(Virtual Port-Channel)，M-LAG该技术只在两台设备的link-agg层面实现二层虚拟化，两台成员设备的管理和控制平面是独立的，M-LAG，支持的跨设备链路捆绑组网，支持Dual-Active的L3GW。在接入侧，从对端设备视角、服务器视角看，M-LAG与堆叠类似。但是，从三层网络角度看，M-LAG的两个成员节点拥有自己独立的IP地址，两个节点有自己独立的管理和控制平面。从架构角度看，MC-LAG的两个成员设备仅存在数据面的耦合，以及协议面的轻量级耦合。



MC-LAG的架构，决定了此技术方案不存在堆叠难解决的三个问题

问题/风险	MC-LAG解决的原因
整系统级可靠性风险	两个成员节点，在软件上只有协议级的弱耦合，因此故障传递的可能性大大降低
版本升级的业务中断时间长	两个成员节点，在软件上只有协议级的弱耦合，支持一个成员节点单独升级。升级期间，另外一台成员负责转发，业务中断时间只有秒级
整系统升级风险	两个成员节点，在软件上只有协议级的弱耦合，支持一个成员节点单独升级。即使升级失败，还有一台成员正常工作。

最后比较堆叠与M-LAG的优缺点，以及场景选择建议

	堆叠	MC-LAG
理论可靠性	较低 管理面、控制面、数据面紧耦合，软件故障可能跨框传递，因此理论可靠性相对较低	高 在软件上只有协议级的弱耦合，因此故障传递的可能性大大降低，理论可靠性高
升级风险	风险高 升级失败将导致堆叠系统下所有业务网络中断，升级风险高	风险低 支持一个成员节点单独升级，即使升级失败，还有一台成员正常工作，升级风险低
升级中断时间	中断时间长 普通升级十分钟级 两框RoundRobin为十秒级~分钟级，和业务相关 ISSU的升级，秒级或者无中断	中断时间短 普通升级秒级 ISSU无中断
网络方案设计	简单 按单节点设计	相对复杂 在三层网络侧需按双节点设计
特性支持度	全面 几乎支持所有特性	相对弱 不支持所有特性，但支持数据中心常用的基础二三层、TRILL、VXLAN等常用协议
配置&维护方便性	简单 只需维护单节点	相对复杂 需要维护双节点

根据上面的对比表格，堆叠和MC-LAG各有优缺点。总的来说，对于网络设计/维护人员，堆叠胜在管理维护简单，MC-LAG胜在可靠性和低升级风险

