

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

CCIE自学认证之

2012

自学Nexus理论与配置



红盟过客

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

作者：ONE 翼

作者：红盟过客 QQ: 119960991

籍贯：湖北宜昌。有宜昌的朋友可以一起聊聊技术。

修正：舍得

[红盟过客内部学习主群 50722586](#) 本群只聊工作经验与生活,进来的人都需要经过测试。

特别欢迎对思科 45, 65,85,76,12410,12416,CRS-1 等高端设备有配置经验的朋友进来相互学习技术。群实行严格管理:只有过硬的管理才有过硬的技术.这是本群的特色---只聊工作经验与生活。生活是指：房子，爱情,理财，学习。工作经验是指：在工作中遇到的问题与经验总结。绝对不聊版本，考试，等等。这是内部学习群。希望大家遵守规则

已完工的作品

《2010 思科路由协议挑战 100 问》 第三期
《2010 思科路由协议挑战 100 问》 第四期
《2011 自学 ccie 语音工程案例与配置》
《2011 自学 ccie 交换工程案例与配置》
《2011 思科路由协议挑战 100 问》 第二期修正
《2009 思科路由协议挑战 100 问》 第一期修正
《2010 思科项目工程师技术经验汇总参考》 第一期
《2010 自学 CCIE 交换工程案例与配置》 第一版
《2009 思科路由协议挑战 100 问》 第三期修正
《2009 mpls vpn 自学命令配置排错》 第一版
《2009 ISIS 命令与配置手册实验笔记》 第一版
《2009 3560 交换机自学命令与配置》 第一版
《2008 vpn 自学命令与配置》 第一版
《2008 ospf 命令与配置手册实验笔记》 第二版
《2008 bgp 命令与配置手册实验笔记》 第二版
《2008 三层交换与路由综合配置案例》
《RHCE5.0 实验手册》
《华为 3600 交换机实验笔记》
《BGP 设计与实现笔记》
《JUNIPER NETSCREEN 防火墙基础实验》
《杭州某中学断网解决办法》

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

目录

第一节 前言.....	5
第二节 个人感谢.....	6
第三节 自励.....	6
第四节 为荣誉而译.....	6
为荣誉而译：人民邮电出版社思科 CCNP 安全认证新书-译者招募啦.....	6
第五节 学习心得.....	7
-----生命的绽放.....	7
第一节 基本概念.....	16
第二节 配置习惯的改变.....	16
第三节 产品介绍 N7000.....	17
第四节 N5000.....	18
第五节 N2000.....	20
第六节 N1000.....	21
第七节 N1000---VN-link.....	22
第八节 N1000---VM ESX Server.....	25
第九节 许可证.....	29
第十节 交换矩阵.....	31
第十一节 M 和 F 系列板卡.....	31
第十二节 电源.....	32
第十三节 VDC.....	33
第十四节 VRF.....	34
第十五节 实验举例.....	34
第十六节 交换.....	37
第十七节 VLAN.....	37
第十八节 端口聚合.....	38
第十九节 MSTP.....	39
第二十节 UDLD.....	39
第二十一节 VPC.....	39
第二十二节 Peer keepalive link.....	41
第二十三节 peer-link.....	42
第二十四节 peer-gateway and switch.....	43
第二十五节 配置 VPC (实验).....	43
第二十六节 FEXes---FEXes 在 NX 系列中的支持.....	45
第二十七节 static pinning.....	46
第二十八节 dynamic pinning.....	47
第二十九节 FEXes 在 NX 系列中的支持.....	48
第三十节 active-active FEX.....	50
第三十一节 N7K 中的 FEX.....	52
第三十二节 FEX 在服务器中的应用举例.....	52
第三十三节 路由.....	55
第三十四节 OSPF.....	56

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

第三十五节 EIGRP.....	56
第三十六节 ISIS.....	57
第三十七节 BGP.....	57
第三十八节 IPV6 基础配置.....	58
第三十九节 组播.....	58
第四十节 --PIM.....	59
第四十一节 IGMP.....	59
第四十二节 MSDP.....	59
第四十三节 管理.....	60
第四十四节 fabric path.....	60
第四十五节 DCNM.....	62
第四十六节 L2, L3 可靠性.....	63
第四十七节 Port-Profile.....	63
第四十八节 ACS.....	64
第四十九节 CFS 部分配置同步协议.....	65
第五十节 ISSU.....	66
第五十一节 ACL.....	66
第五十二节 QOS.....	66
第五十三节 CMP.....	67
第五十四节 BFD.....	67
第五十五节 配置回滚.....	67
第五十六节 FCOE.....	68
第五十七节 SCSI、FC、FCOE 概念.....	68
第五十八节 SCSI\FC 传输机制.....	69
第五十九节 FCOE 机制和配置举例.....	70
第六十节 FCOE 机制的要点.....	73
第六十一节 个人推荐的 FCOE TOP.....	74

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

第一节 前言

记得和老代（红盟过客）的认识是在 2007 年了，那个时候应该算是群里的第一批吧。

当时老代已经小有名气，不过当时作品还不是特别多，到后来有幸参与了 100 问中实际工程的那一期，此后由于各种原因，没再参与太多。

老代群里各路高手云集，具体名字就不多说，经常在群的也都心里有数。在老代的带动下，大家都十分具有共享精神，各路资料汇集于此，鄙人虽然一直在业界工作，但也确实受益匪浅。

非常感谢各位兄弟无私的共享自己的资料和心得，在这里同样将我自己的一份力量献给大家。在本次编写中，鄙人笔记非常凌乱，也有很多漏洞，全依赖老代的辛勤整理和补漏才有能发布的样子，呵呵。s

特别申明：

所涉及的特性和资料都是建立于 NX-OS 5.0 的基础上，对以后 OS 更新后的特性不保证，本手册中如有错误欢迎大家指出，但是本手册目的单纯只是作为参考资料，非官方文档，所以请大家尽量最终向思科原厂工程师确认，谢谢！

培训心得：

为期一整周的原厂培训，上课的老师是一位广东人，普通话我适应了很久，特别是在学习新技术方面存在一定的难度，所以在最后一天的 FCOE 部分笔记做的不是很好，经过后期的查阅资料进行了补充。

虽然有实验环境，但是为远程，实际设备不在上海，而且登录也是 VDC，不是一个人独享一台 N7K，要配置一个完整的实验，要 4 个人协作，呵呵，本来课后时间就非常短，基本只有半个小时，5 点准时被 clear，要给国外的朋友。所以基本就上去逛了逛，熟悉了一下界面和各种命令，结合思科官方的手册，模拟出来的实验 TOP 和配置。

思科的培训场地比较到位，有无线接入以及足够多的桌子和接线板，所以也就有了这份笔记，如果单纯手写，大部分人已经不太会写了吧，呵呵，而且也确实没有笔记本打字来的快。讲师水平绝对足够高，基本全方位的知识都讲的不错，而且深度绝对够，能给思科原厂工程师培训的讲师不会太差。

但是本次来培训的人也都不是省油的灯，因为思科前期反复要求过，呵呵。大家得着问题就猛问，特别几个思科原厂工程师问的更是犀利，大家大有打破沙锅问到底的精神，不屈不饶，最后终于把讲师问倒了几次，个人感觉谁要是一年里面天天参加这种培训，出来一定是大牛中的大牛。

这本手册里面很有部分内容都是我们问出来的，比如 SUP1 引擎不负责数据转发等等。

结语：

没有我的公司就没有这份手册，所以首先感谢我的公司，所有权第一是我现在的公司，第二是老代和群里的各位朋友，感谢你们！

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

第二节 个人感谢

仅以此作品献给最伟大的母亲和为了理想支持我实验室的兄弟

Love means never having to say you're sorry.爱---就是永远不必说对不起。(《爱情故事》)

首先说明：这份文档是杭州 one 翼这位兄弟思科培训的时候记的笔记。其实我什么工作也没有做，只是把它发过来的整理成文档罢了。我自己认真的看了这份文档，很有信心的讲。这份文档是目前关于 nexus 最完整的文档。

one 翼最早是在与我合写 100 问的时候，提出了 30 多个实际工程案例。他的工程经验相当的丰富，就像他自己讲的，不管是我的文档还是在我们项目工程 TAC 群里。确实学习到了很多的经验和知识，所以将自己的宝贝拿出来与大家一起学习。

红盟过客工程 TAC 《主》 50722586 群的唯一目地：帮助大家解决在工程项目中遇到的问题，希望更多有经验与能力的朋友加入。

语音的文档好了，那些曾经只要拍过我的作品的朋友要记得找我要，一定要找我要了。

项目排错案例，由于案例太多，本想只出一份文档的，将近 300 个案例，没有办法，准备出三份文档。一期差不多在八月份完工。

我自己的 taobao <<http://ccie19.taobao.com/>>>将在明年年底前关闭，设备将分批自己购买。感谢那些不知道我能否成功依然坚持支持我的朋友。正是由于你们的支持，我才有成功的可能。

感谢杭州的王嘉啸，没有你的帮助,我将一事无成。

十年磨一剑，成功的背后是漫长辛勤的付出。

知识需要长期辛勤地学习钻研，才能领悟其精华；方法需要反复地实践，才能得心应手；经验需要不断地总结，才能厚积薄发；心理需要精心地调理，才能日趋成熟。

第三节自励

Stay hungry,stay foolish.求知若饥，虚心若愚--乔布斯

不管是生活还是工作，自己也经常遇到打击与挫折。

七月至少有一个月有没有休息好了，时常在半夜里醒来，因为自己负责的一个项目一直没有签定下来。其实年轻的时没有失败，还是那句俗语：少壮不努力，老大徒伤悲。

生命是一个奋斗的过程，也是一个等待的过程。因为人生不会总是一马平川，不会总是春风得意。有太多的不顺心、不如意甚至挫折沮丧面前，我们需要的是平和的心态，在等待与努力中积聚力量，最后实现灿烂的绽放。

第四节 为荣誉而译

为荣誉而译：人民邮电出版社思科 CCNP 安全认证新书-译者招募啦

<http://bbs.56cto.com/thread-81313-1-1.html>

人邮的傅工现在手上有四本新书考试认证书目：新书每本厚大约在 700 到 800 页。英语原版本的大约在七月底才会上市。这是考试用书。时间限定在四个月内译完。

本作品仅赠于本档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

可以直接联系我

时间是把杀猪刀，刀刀催人老。假如回到过去，能做点什么让现在的生活更有意义呢能为人邮翻译新书，是我们每个做网络工程师莫大的荣誉。

至少十年以后，当自己不在年轻的时候，想到年轻时，曾经在人邮的书上留下自己的名子,是一件多令人自豪的事情。

当然，如果单纯为了赚钱，请放弃，因为一本书的译书工作可能要好几个月。每本书的译费也只在 1.5 万左右。

人邮对译者有严格的要求，所以，试译后，没有被录取，也是件很正常的事情，因为这是对译者，责任心，态度，语言功底，技术要求的一次全面考验。

所以，为荣誉而译。

我们 56cto 的全体同仁将全力支持为人邮的工作。

Security IPS 642-627 Official Cert Guide 700 页

Security VPN 642-647 Official Cert Guide 800 页

Security Firewall 642-617 Official Cert Guide 700 页

Security Secure 642-637 Official Cert Guide 800 页

我相信：在译者田果，孙余强，等一些不管是责任心还是技术要求都给大家做榜样的基础上，人邮一定会给大定带来，更多更好，更高品质的思科专业书籍。

第五节 学习心得

-----生命的绽放

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。摘自《珍珠港》

生活中，我们常常遇到不公平或歧视。弱者，只会徒劳地愤愤，喋喋地牢骚。而聪明的你，却将它化成一股奋进的动力，鞭策磨砺自己，努力、努力、再努力，证明自己，实现梦想。

这二个月认识了几位新的朋友。有《Cisco 局域网交换机配置手册(第 2 版)》的译者付强。还有人邮的傅道坤,就像他自己讲的，大家都以为他是一个很年长的朋友，原来也是 80 后的，与我只相差一岁罢了。

还认识了路由结构的译者孙余强。孙工是一个极其负责的译者。他译的路由结构这本书我在我们的论坛 56cto 上面重点推荐，希望喜欢 bgp 的朋友一定要认真的看看这本书。我自己一个月之内差不多读了三遍了。虽然读第一遍的时候有一些看不太明白，但是读到第三遍的时候差不多就没有遇到什么问题。每个人都不是神，不可能看书一遍就能看懂和能记住的，像卷一这样的书，自己差不多每二个月都会重看一遍。温故而知新。

付强这位朋友可以说是年轻有为，现在正当是壮年，年轻的时候应当为自己的梦想而努力，虽然交换这本书是第一次译，但是，自己也是认真的看了好几遍，找不到硬伤。当然，里面译的也还有一些不太正确的地方。

译者邮的专业书绝对不是一件轻松的事情，要想译好更是不简单的。虽然他是第一次译，但是质量还是相当的不错。我给好几位喜欢交换命令的朋友推荐了这本书。当然，他现在还在译路由的配置手册，到时候，如果时间及精力来得及，我也会去做做校正的工作。

我与傅工谈了很长时间。都是关于人邮的思科书。人邮的思科书我差不多都有的。当然，

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

我也可以很明显的感觉到,自从他进入到人邮以后,由他负责的技术书很明显的质量提高很多。特别是加入了很多的译者注,让读者理解起来不在麻烦与难懂。人邮以前的一些书,说实话,译的质量只能说一般般。特别是有几位译的 EIGRP,我只能说,译出这样的水平,害了很多读者。不光花钱买了书,并且会误人子弟的。

还认识了田果这位朋友。他是 ASA 与网络安全解决方案的译者。很明显,他译的书质量也是相当的有保证。自己的 ASA 这一块很差,所以,他译的 ASA 指南自己是第一时间买回来,认真的学习与实验。技术有时候更新太快,当大家在玩 8.2 的时候,8.42 又出来了。并且 8.42 有了很多更新。

虽然与田果朋友素位某面,但是有一种英雄相惜的感觉。希望他能给我们带来更多更好的译书。

这一段时间,关于思科的负面消息不断。其实是很正常的事情,并且有一个问题,大家有没有想过,思科有可能会倒闭,我想,这种可能性是绝对有的,在 08 年,美国有 100 多年的投行都能倒闭,思科有什么不可能的。所以大家一定要认清楚,思科的 ccie 只是一个厂家的认证,要轻认证,重技术。并且,我个人感觉到思科有极强的霸权主义思想,比如。人邮,引进了思科的图书,就不能引进其它的图书,所以大家在市面上看不到人邮出版的 XXXX (或 juniper、或华为) 图书,原因就在这里。

所以我很早就发现了思科的这样的美式霸权主义思想。但是,思科也是一种很严谨的厂家。宁可将二台 65 进行 vss 堆叠,也不愿意虚报数据。现在思科最新的 65 的引擎是 sup 2T。如果国产的设备能有思科这样一半的严谨的态度与管理方式,何愁打不败思科这样的厂商。希望中国的设备能越来越真实。

本打算 100 问 5 期写关于 mpls 的技术的。但是后来决定写一下关于思科新产品的文档好了。有喜欢思科售前的朋友也可以一起来写写。

写文档不为名,不为利,只为兴趣。如果有那一天我自己不在写思科的文档了,那就说明我终于将自己喜欢的思科给放弃了。自己是一个很讲原则的人,是就是。不是就不是。

我是 56cto 的管理员,有时候也会想到,如何将 56cto 做的更好,更加强大。我个人感觉虽然 56cto 在认证等方面确实做的很不错。但是在一些方面也做的不够好。

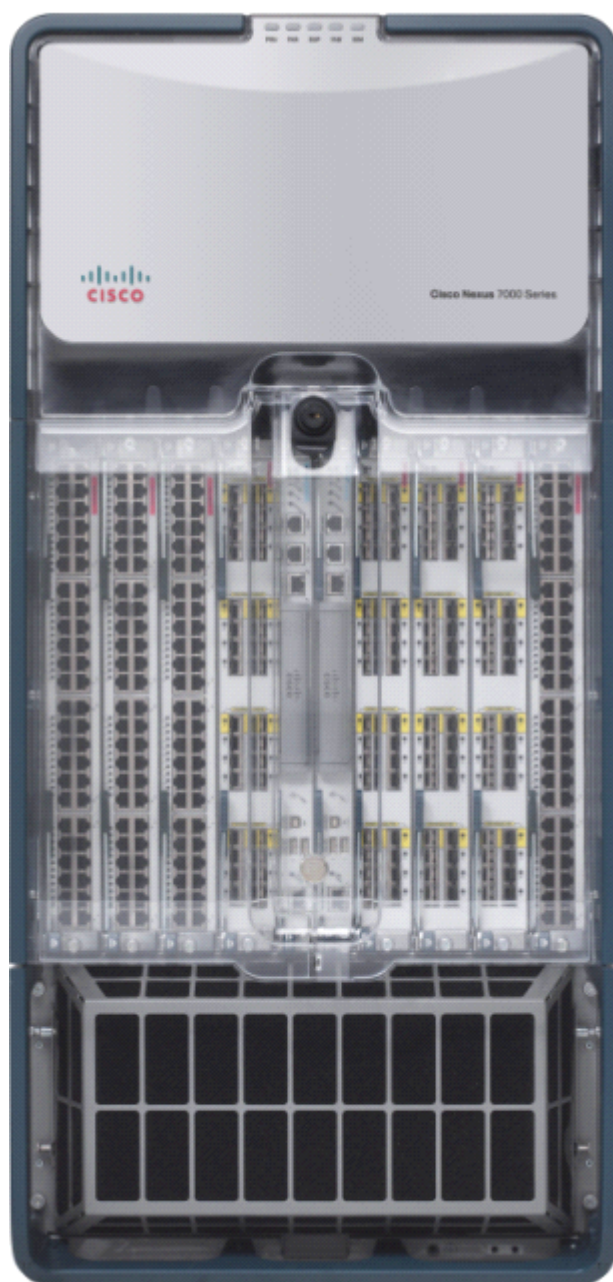
当然,我们 56cto 也会大力加强与人邮的一些合作。可以为入邮的新书好书做一些好的宣传。也算是贡献自己的一份力量罢了。

代工写于 2011 年 8 月 2 日于杭州

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



7010

本作品仅赠于本文档所有付出与淘宝的朋友
<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

Product	Description	Quantity	Price
			63,000.
N7K-C7018	18 Slot Chassis, No Power Supplies, Fans Included	1	00
N7KS1K9-51	Cisco NX-OS Release 5.1	1	0
			31,500.
N7K-LAN1K9	Nexus 7000 LAN Enterprise License (L3 protocols)	1	00
	Nexus 7000 - 48 Port 10/100/1000, RJ-45, 40G		31,500.
N7K-M148GT-11	Fabric	1	00
			56,700.
N7K-M148GS-11	Nexus 7000 - 48 Port 1G, SFP, 40G Fabric	1	00
			1,493.0
GLC-LH-SM	GE SFP, LC connector LX/LH transceiver	1	0
GLC-SX-MM	GE SFP, LC connector SX transceiver	1	750
GLC-T	1000BASE-T SFP	1	830
			8,390.0
GLC-ZX-SM	1000BASE-ZX SFP	1	0
			2,310.0
SFP-GE-L	1000BASE-LX/LH SFP (DOM)	1	0
			1,155.0
SFP-GE-S	1000BASE-SX SFP (DOM)	1	0
SFP-GE-T	1000BASE-T SFP (NEBS 3 ESD)	1	924
			8,390.0
SFP-GE-Z	1000BASE-ZX Gigabit Ethernet SFP (DOM)	1	0
	Nexus 7000 - 32 Port 10GbE, 80G Fabric (req. SFP+)		147,000
N7K-M132XP-12		1	.00
			21,000.
SFP-10G-ER	10GBASE-ER SFP Module	1	00
			8,390.0
SFP-10G-LR	10GBASE-LR SFP Module	1	0
			3,140.0
SFP-10G-SR	10GBASE-SR SFP Module	1	0
N7K-M148GS-11	Nexus 7000 - 48 Port GE Module with XL Option		56,700.
L	(req. SFP)	1	00
			1,493.0
GLC-LH-SM	GE SFP, LC connector LX/LH transceiver	1	0
GLC-SX-MM	GE SFP, LC connector SX transceiver	1	750
GLC-T	1000BASE-T SFP	1	830
			8,390.0
GLC-ZX-SM	1000BASE-ZX SFP	1	0
			2,310.0
SFP-GE-L	1000BASE-LX/LH SFP (DOM)	1	0
SFP-GE-S	1000BASE-SX SFP (DOM)	1	1,155.0

本作品仅赠于本档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

			0
SFP-GE-T	1000BASE-T SFP (NEBS 3 ESD)	1	924
			8,390.0
SFP-GE-Z	1000BASE-ZX Gigabit Ethernet SFP (DOM)	1	0
N7K-M108X2-12			92,400.
L	Nexus 7000 - 8 Port 10GbE with XL option (req. X2)	1	00
			1,260.0
X2-10GB-CX4	10GBASE-CX4 X2 Module	1	0
			21,000.
X2-10GB-ER	10GBASE-ER X2 Module	1	00
			8,400.0
X2-10GB-LR	10GBASE-LR X2 Module	1	0
			3,140.0
X2-10GB-LRM	10GBASE-LRM X2 Module	1	0
			6,290.0
X2-10GB-LX4	10GBASE-LX4 X2 Module	1	0
			4,190.0
X2-10GB-SR	10GBASE-SR X2 Module	1	0
			33,600.
X2-10GB-ZR	10GBASE-ZR X2 Module	1	00
N7K-F132XP-15	Nexus 7000 - 32 Port 1G/10G Ethernet Module, SFP/SFP+	1	105,000
			.00
			1,493.0
GLC-LH-SM	GE SFP, LC connector LX/LH transceiver	1	0
GLC-SX-MM	GE SFP, LC connector SX transceiver	1	750
GLC-T	1000BASE-T SFP	1	830
			8,390.0
GLC-ZX-SM	1000BASE-ZX SFP	1	0
			8,390.0
SFP-10G-LR	10GBASE-LR SFP Module	1	0
			2,720.0
SFP-10G-LRM	10GBASE-LRM SFP Module	1	0
			3,140.0
SFP-10G-SR	10GBASE-SR SFP Module	1	0
			2,310.0
SFP-GE-L	1000BASE-LX/LH SFP (DOM)	1	0
			1,155.0
SFP-GE-S	1000BASE-SX SFP (DOM)	1	0
SFP-GE-T	1000BASE-T SFP (NEBS 3 ESD)	1	924
			8,390.0
SFP-GE-Z	1000BASE-ZX Gigabit Ethernet SFP (DOM)	1	0
SFP-H10GB-ACU			
10M	Active Twinax cable assembly, 10m	1	861

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

SFP-H10GB-ACU				
7M	Active Twinax cable assembly, 7m	1	756	
SFP-H10GB-CU1				
M	10GBASE-CU SFP+ Cable 1 Meter	1	315	
SFP-H10GB-CU3				
M	10GBASE-CU SFP+ Cable 3 Meter	1	441	
SFP-H10GB-CU5				
M	10GBASE-CU SFP+ Cable 5 Meter	1	546	
N7K-M148GT-11	Nexus 7000 - 48 Port 10/100/1000 Module with XL		31,500.	
L	option	1	00	
N7K-M132XP-12	Nexus 7000 - 32 Port 10GbE with XL Option, 80G		147,000	
L	Fabric (req.	1	.00	
			21,000.	
SFP-10G-ER	10GBASE-ER SFP Module	1	00	
			8,390.0	
SFP-10G-LR	10GBASE-LR SFP Module	1	0	
			2,720.0	
SFP-10G-LRM	10GBASE-LRM SFP Module	1	0	
			3,140.0	
SFP-10G-SR	10GBASE-SR SFP Module	1	0	
SFP-H10GB-ACU				
10M	Active Twinax cable assembly, 10m	1	861	
SFP-H10GB-ACU				
7M	Active Twinax cable assembly, 7m	1	756	
SFP-H10GB-CU1				
M	10GBASE-CU SFP+ Cable 1 Meter	1	315	
SFP-H10GB-CU3				
M	10GBASE-CU SFP+ Cable 3 Meter	1	441	
SFP-H10GB-CU5				
M	10GBASE-CU SFP+ Cable 5 Meter	1	546	
	Nexus 7000 - Supervisor, Includes External 8GB Log		52,500.	
N7K-SUP1	Flash	1	00	
	Nexus Compact Flash Memory 2GB (Expansion		2,520.0	
N7K-CPF-2GB	Flash - Slot 0)	1	0	
	Nexus 7000 - Supervisor, Includes External 8GB Log		52,500.	
N7K-SUP1	Flash	1	00	
	Nexus Compact Flash Memory 2GB (Expansion		2,520.0	
N7K-CPF-2GB	Flash - Slot 0)	1	0	
N7K-C7018-FAB-	Nexus 7000 - 18 Slot Chassis - 46Gbps/Slot Fabric		63,000.	
1	Module	5	00	
N7K-AC-7.5KW-I	Nexus 7000 - 7.5KW AC Power Supply Module		63,000.	
NT	International (cab	4	00	
N7K-C7018-FD-	Nexus 7018 Front Door Kit	1	3,150.0	

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

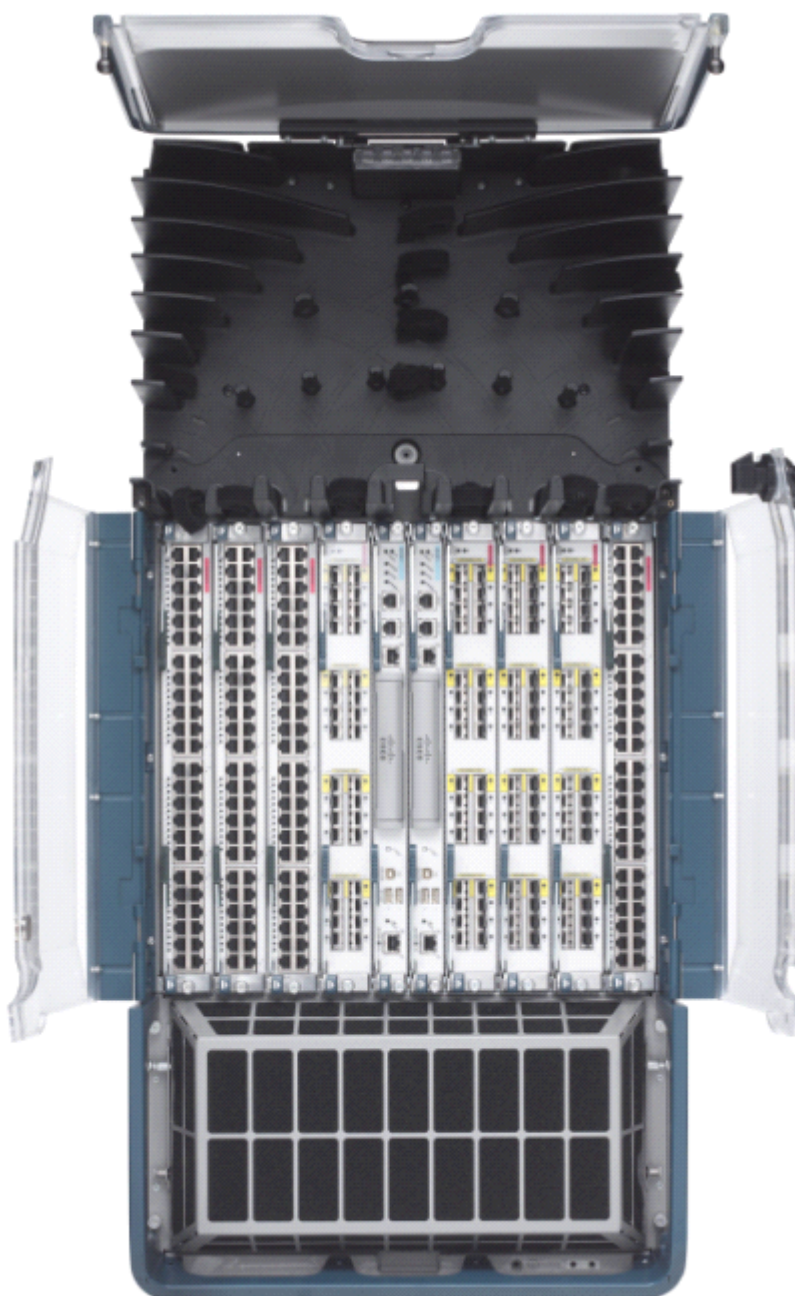
victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

MB			0
N7K-SUP1-8GBU			
PG	Nexus 7000 Supervisor 1 8GB Memory Upgrade Kit	1	0
N7K-SUP1-8GBU			
PG	Nexus 7000 Supervisor 1 8GB Memory Upgrade Kit	1	0
N7K-C7018-FAN	Nexus 7000 - 18 Slot Fan	2	0
CON-OSP-N7KC	ONSITE 24X7X4 18 Slot Chassis, No Power Supplies,		3,600.0
7018	Fans	1	0
CON-OSP-C701	ONSITE 24X7X4 Nexus 7000 - 18 Slot Slot		3,600.0
8FB1	Chassis-46Gpbs	5	0
CON-OSP-7F32	ONSITE 24X7X4 Nexus 7000 - 32 Port 1G/10G		2,625.0
XP	Ethernet Fab	1	0
CON-OSP-N7LA			
N	ONSITE 24X7X4 Nexus 7000 LAN Enterprise Lic	1	900
CON-OSP-N708			3,300.0
X2L	ONSITE 24X7X4 Nexus 7000 - 8 Port 10GbE	1	0
CON-OSP-N732	ONSITE 24X7X4 Nexus 7000 - 32 Port 10GbE, 80G		5,252.0
XP	Fabric	1	0
CON-OSP-7M32	ONSITE 24X7X4 Nexus 7000 - 32 Port 10GbE with XL		5,252.0
XPL	Opt	1	0
CON-OSP-M148	ONSITE 24X7X4 Nexus 7000 - 48 Port 1G, SFP, 40G		2,025.0
GS11	Fabric	1	0
CON-OSP-N748			2,025.0
GSL	ONSITE 24X7X4 Nexus 7000 - 48 Port GE Module	1	0
CON-OSP-N748	ONSITE 24X7X4 Nexus 7000 - 48 Port 10/100/1000,		1,125.0
G	RJ-45	1	0
CON-OSP-N748	ONSITE 24X7X4 Nexus 7000 - 48 Port 10/100/1000		1,125.0
GL	Module	1	0
CON-OSP-N7SU	ONSITE 24X7X4 Nexus 7000 - Supervisor, Includes		3,000.0
P1	Ext	1	0
CON-OSP-N7SU	ONSITE 24X7X4 Nexus 7000 - Supervisor, Includes		3,000.0
P1	Ext	1	0
			127679
			3

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



Product	Description	Quantity	Price
N5K-C5020P-B F	N5000 2RU Chassis no PS 5 Fan Modules 40 ports (req SFP+)	1	72,450.00
N5KUK9-421N 1.1	Nexus 5000 Base OS Software Rel 4.2(1)N1(1)	1	0
N5000FMS1K9	Nexus 5000 Fabric Manager Server License	1	14,700.

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

			00
			50,400.
N5020-SSK9	Nexus 5020 Storage Protocols Services License	1	00
	N5000 1000 Series Mod 4x10GE 4xFC 4/2/1G(req		10,920.
N5K-M1404	SFP+/SFP)	1	00
			6,720.0
N5K-M1060	N5000 1000 Series Module 6port 8/4/2/1G FC	1	0
N5K-PAC-1200			6,300.0
W	Nexus 5020 PSU module, 100-240VAC 1200W	2	0
SFS-250V-10A-			
CN	SFS Power Cord - 250V, 10A - PRC	2	0
N5020-ACC-KI			
T	Nexus 5020 Accessory Kit, Option	1	0
CON-OSP-N5S	ONSITE 24X7X4 Nexus 5020 Storage Protocols Svc		5,400.0
SK	License	1	0
CON-OSP-N50			8,280.0
20	ONSITE 24X7X4 N5000 2RU Chassis no PS 5	1	0
CON-OSP-N50			
60	ONSITE 24X7X4 N5000 1000 Series Module	1	768
CON-OSP-N54	ONSITE 24X7X4 N5000 1000 Series Mod 4x10GE 4xFC		1,248.0
04	4/2/1G	1	0

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

第一节 基本概念

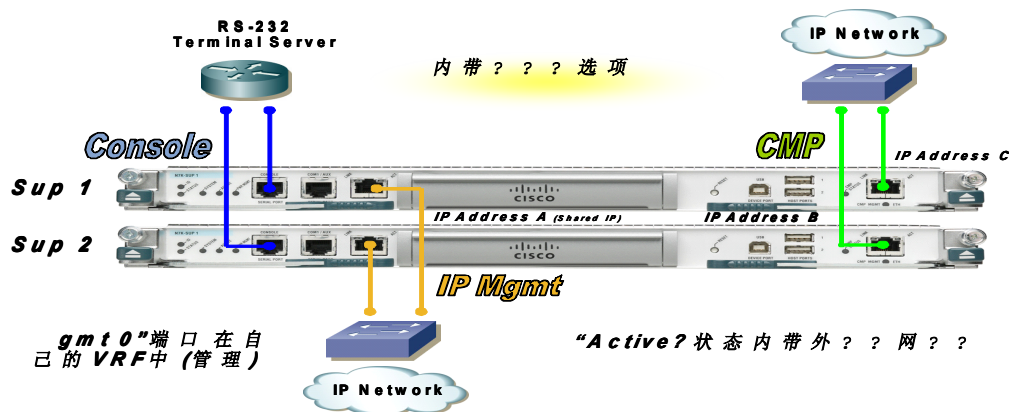
思科官方的各数据中心互联带宽PPT上默认画的都是10G

1Tier: N7K接2K, 等于将N7K一个10G口变成多个1G口。L2交换还是由N7K处理。

2Tier: N7K-N5K-N2K 以及N7K-N7K-MDS,做port-channel最好不要跨板块, 要么都用M做, 要么都用F做。(实验室实际操作也是不可以) Nexus系列里, 所有端口不分10GE和GE都叫 ethernet X/X,

N7K和N5K都有各自很多LIC, 要根据情况分别购买, 比如N5K具有: 高级版LIC, 支持光存储版LIC等等。如果是后期买LIC, 要提供很多资料。

所有的VDC共用SUP1引擎上的管理口, 一个物理端口, 多个逻辑地址, 不同的地址对应不同的VDC管理登陆。



第二节 配置习惯的改变

interface eth 2/1-3 或者inter eth 2/1,eth 2/4-6 代替inter range

举例: N7K-1-pod3(config)# int e1/18,e1/20,e1/22,e1/24

config 代替config t, 不再使用do show 而是直接show

某些命令, 不离开类似config-if模式的时候, 无法生效了。

配置trunk只需要: switch mode trunk 一句即可 (ISL已经取消)

NX-OS系列所有的特性都要手动先启动。

UDLD: 单向连接侦测, N7K启用 (featrue udld)之后, 所有10G自动执行UDLD, 1G的铜缆需要手动。

NX-OS系列新增参数组的功能, 将一组命令组成一个 profile,将这个pro套用于指定的接口。profile之间可以嵌套, 一个profile可以采用字母或者数字来命名。

NX-OS系列不支持DTP, 必须手动配置指定的静态trunk, TRUNK仍然默认允许所有VLAN通过。两边对联设备本真VLAN必须一致。

NX-OS, 所有VDC VLAN之和不能超过16000个。也就是4个4000

NX-OS现在只支持LACP, 不支持PAGP了。而且只能静态手动配置了。

NX-OS 根据对端STP切换到原老的STP协议。实际应用中, 却有一些问题。但是如果对端老

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

设备也配置支持RSTP或者PVST，就没有问题。

span port type edge 代替 span portfast

NNX-OS中只有类似IOS中的扩展ACL，掩码可以使用/24之类的数字代替。

第三节 产品介绍 N7000

N7K系列仍然不支持任何服务模块，必须得与C6500结合。

将安装了FWSM或者是ACE模块的6509以旁路模式部署在N7K上，等于是将带FWSM的6509当作了N7K的服务模块。



N7010



N7018

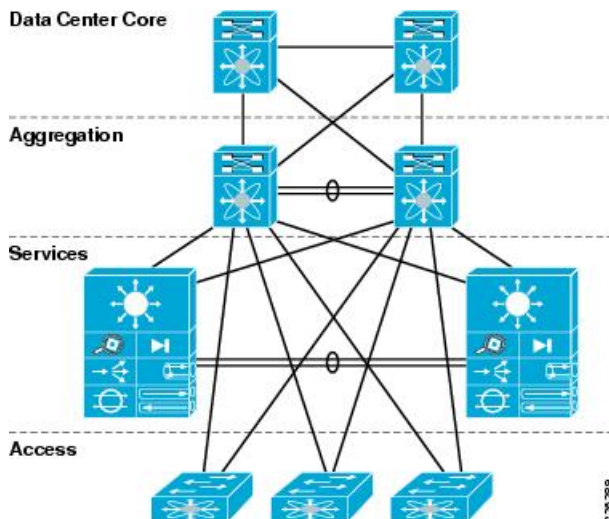
N7K中，所有VDC VLAN之和不能超过16000个。也就是4个4000

N7K连接N2K，必须使用以太通道方式互联，强制的，不支持单条普通线路（而N5K可以）。

N7K可以使用hardware命令限制各种流数据包抵达主控板的数率，防止主控板死机。

N7K也可以提供链路层加密。

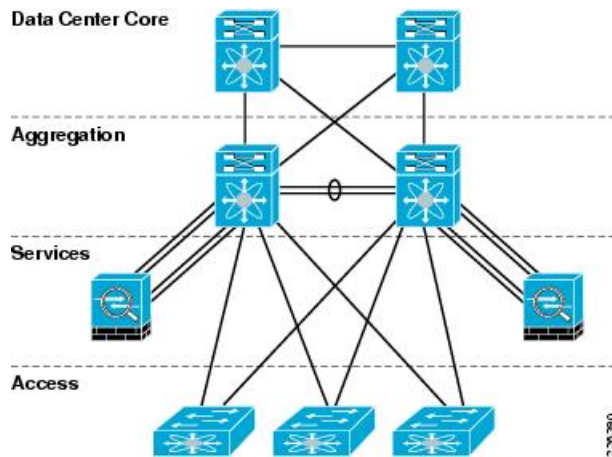
N7K 不支持各种服务模块，以下为在DC中，使用CISCO各服务模块示意图：



本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



第四节 N5000

N5500支持三层，N5000只能二层

5010默认带20个 10E/FC0E/DCB口 1个扩展槽 5020默认40个 2个扩展



现在新出一种模块，可提供8个口，或者是以太光口（也可以配为FCOE），或者是存储的FC光口。



Ethernet
6-Port 10G/FCoE/Data
Center Ethernet



Ethernet + Fibre Channel
4-Port 10G/FCoE/Data
Center Ethernet
4-Port 1/2/4G Fibre Channel

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



Fibre Channel
8-Port 1/2/4G Fibre Channel

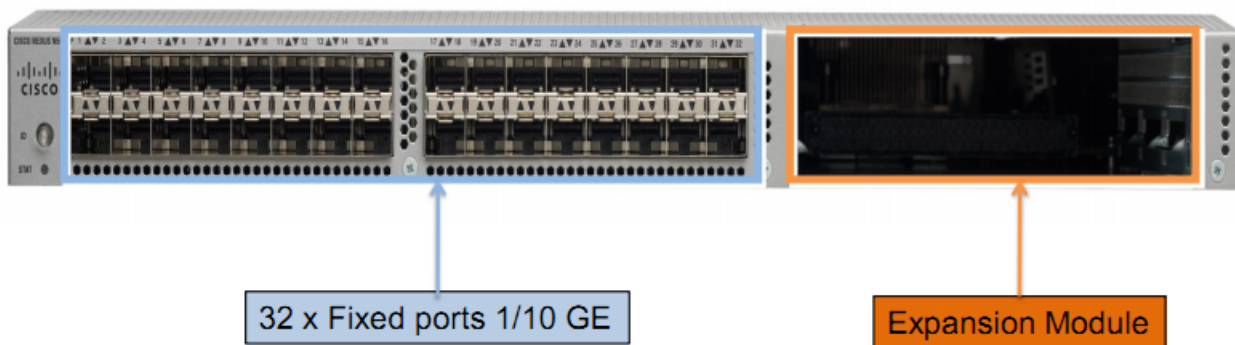


Fibre Channel
(Q3CY09)
2/4/8G Fibre Channel

N2000, 在N5K上启用FEX 延伸特性, 并定义FEX 号码, 并指定为FEX模式, 用于连接N2K 每一台N2K对应自己的唯一FEX的号码。



5548:



本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

N5596 96x SFP+ (w/3x Module) in 2RU



96 Fixed Ports capable of 10GE/1GE/FCoE/FC (Unified Port)

Three Expansion Module Slots

Modules

16x 10G/FCoE/DCB SFP+

8x 10G/FCoE, 8x 1/2/4/8G FC

16x Unified Ports SFP+

Supports on Nexus 5500 Platform features AND unified ports on all ports

第五节 N2000

N2000不是独立的交换机，本身没有交换功能，同一个N2K中，两个相互端口互访，需要上升到N5K去交换（性能不会因此降低，可以简单的看作N5K的物理端口扩展而已）N2000K，在N5K上启用FEX 延伸特性，并定义FEX 号码，并指定为FEX模式，用于连接N2K每一台N2K对应自己的唯一FEX的号码。

两台N5K看同一个N2K的时候，各方面完全一致。针对N2K的各种配置也要对应一致。

N7K连接N2K，必须使用以太通道方式互联，强制的，不支持单条普通线路（而N5K可以）。



本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

Nexus 2148T, 2248TP, 2232PP Capabilities

Features	2148	2248	2232
Fabric Ports (NIF)	4	4	8
Fabric Link Port Speed (NIF)	10Gbps	10Gbps	10Gbps
Port Channels on Fabric Links (NIF)	1 x 4 ports maximum Hash L2/L3 fields	1 x 4 ports maximum Hash L2/L3/L4	1 x 8 ports maximum Hash L2/L3/L4
Host Ports (HIF)	48	48	32
Host Port Speeds (HIF)	1 Gbps only	100Mb/1Gbps	1Gbps/10Gbps (No 1 Gbps at FCS, 2HCY10)
Local Port Channels on Host Ports (HIF)	Not Supported	Max 8 ports per port channel, Max of 24 port channels per 2248 Hash L2/L3/L4	Max 8 ports per port channel, Max of 16 port channels per 2232 Hash L2/L3/L4
FCoE/DCB Servers	No	No	Yes
Supported Parent Switches	Nexus 5010/5020/5548 Nexus 5596 - Q1CY11	Nexus 5010/5020/5548 Nexus 7000 M1 Nexus 5596 - Q1CY11 Nexus 7000 F2 - 2HCY11	Nexus 5010/5020/5548 Nexus 5596 - Q1CY11 Nexus 7000 M1 - Q1CY11 Nexus 7000 F2 - 2HCY11

Description	Specification
Cisco Nexus 2148T	48 1000BASE-T host interfaces and 4 10 Gigabit Ethernet fabric interfaces (Small Form-Factor Pluggable Plus [SFP+])
Cisco Nexus 2224TP	24 100/1000BASE-T host interfaces and 2 10 Gigabit Ethernet fabric interfaces (SFP+)
Cisco Nexus 2248TP	48 100/1000BASE-T host interfaces and 4 10 Gigabit Ethernet fabric interfaces (SFP+)
Cisco Nexus 2232PP	32 1/10 Gigabit Ethernet and Fibre Channel over Ethernet (FCoE) host interfaces (SFP+) and 8 10 Gigabit Ethernet and FCoE fabric interfaces (SFP+)
Cisco Nexus 2232TM	32 1/10GBASE-T host interfaces and modular uplinks (8 10 Gigabit Ethernet fabric interfaces (SFP+))

第六节 N1000

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

N1000系列有两个成员： N1000软件，N1010硬件，两者的关系很像软件版本的ACS和1U硬体的ACS

两个模块：

VEM：虚拟以太网模块，是单个物理机器中的虚拟线卡，线卡中，每个虚拟接口接一台虚拟服务器。

VSM：虚拟管理模块，可以在VM或者独立物理机安装执行。一个VSM管理多个VEM，你可以看做是一个虚拟的主引擎，而且还是双引擎，主备模式。

N1000：

N1000V是软件，并不是物理SW，将N1000V安装在一台物理服务器中，CISCO和VM合作将原内置的SW转为新内置的1000SW虚拟交换机，每一台虚拟服务器都连接在本台N1000V的虚拟交换机的一个虚拟端口中。

这样配置之后的服务器称为 CISCO VEM通过N1010，所有的VEM成为N1010中的一个板卡（一个VEM对应一个板卡），由V1010统一管理。

N1010中安装有VSM，来管理众多的VEM

N1000V可以独立运行，也可以受N1010管理，有点类似WLC和AP的概念

N1000V和N1010这个系列，有两种模式，其中一次可以独立于NEXUS之外存在，但是与NK-OS系列结合将取得更好的数据中心特性，比如在N5K和N2K中采用VPC架构与服务器双网卡互联时，能更体现优势。

目前有两种模式可供选择，其中一种模式不要求N1000一定要与NX系列设备联动，但是另一种是必须和N5K结合使用（不再需要N1000V的 VSM）。两种模式都用到VN-link,不过与物理机建立VN-LINK的对象不同而已。

一台物理机器中，如果是N1KV模式，那么同VLAN的不同VM之间的访问本地转发，不同VLAN需要上升到物理交换机处理，所以尽量将同一台物理机器中的VM放在同一个VLAN。如果是N1V（必须结合物理交换机）模式，那么所有的数据都是上升到物理交换机处理，这种模式实际上将所有的交换功能提升到支持VN-LINK的物理交换机中进行处理。

在N1000V模式下，如果物理服务器中的VM肯定都属于一个VLAN，那么也可以将物理交换机和网卡之间配置成为access模式，如果会有多个VLAN，那么建议trunk。

第七节 N1000---VN-link

vn-link:这种模式将使用N5K取代N1K的服务端软件，每个物理机中的VM，好像连接到N5000中的一个口，通过N5K来进行硬件转发。

Vn-link的作用就是将虚拟机对应的虚拟网卡（Vnic）和CISCO交换机（具有VN-LINK功能的）之间建立一个逻辑的链路。这就好比是服务器(虚拟机)的网卡和接入交换机之间用一条线连起来。实现VN-link 有两种途径，一种是通过nexus 1000v来实现，另一种叫NIV(network interface virtualization)是通过硬件交换机（带VN-LINK功能）配合特殊网卡（带interface virtualizer功能，比如思科的polo卡）来实现。

一、通过nexus 1000V实现VN-LINK

VMware提供了第三方在母机上布署软件交换机的API接口。Nexus 1000v 交换机其功能主要是建立一些ACCESS接口（功能NIV件环境下的vEth口）用于连接虚拟机，和一些uplink口，uplink口对应到物理网卡，和上游真实网络交换机相联（该口一般为trunk口，因为要

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

为多个ACCESS口所接的虚拟机做数据转发，这些虚拟机有可能不在同一个vlan)。

1. Nexus1000V由以下两部分组成：

1) Virtual Ethernet module (VEM)-data plane -----Each hypervisor is embedded with one VEM

2) Virtual supervisor module (VSM)-control plane----独立装在一个物理或虚拟的服务器上(可以和vCenter装在一台服务器上)用来管理所有的Nexus 1000V system(包括VSM自己)。一个VSM可以管理64个VEM，VSM也可以工作在Active/Standby的HA模式。VSM负责运行控制层面协议以及管理VEM，但是不会参与数据转发。

VEM负责在同一个母机中各虚拟机之间的数据转发(数据不出母机)，同时还负责将去往其他母机的数据转发到物理的接入层交换机。

2. 在NEXUS1000V中有两种交换接口：

1) Eth指的是母机上的NIC，比如(config) #int eth2/1，这里的模块号是ESX作为VEM加入nexus 1000时被分配的。接口号则跟网卡的先后顺序号有关。

2) vEth就是N1K中的ACCESS口，这种口完全是虚拟的，是通过命令来建立的，为了方便vMotion这种口是没有模块号的，也就是说这种口是和具体某台母机(ESX)没有关系的，可以在所有母机中移动，具体会停留在哪个母机中由对应的虚拟机在哪台母机中决定(其实是一个vEth pool)。

1. NEXUS 1000V VN-LINK实现过程

1) 建立profile

```
(config)# port-profile webserver
```

```
(config-port-prof)# switchport access vlan 10
```

```
(config-port-prof)# ip access-group 500 in
```

```
(config-port-prof)# inherit port-profile server
```

The port profile can then be assigned to a given vEth interface as follows:

```
(config)# interface veth1
```

```
(config-if)# inherit port-profile webserver
```

先建profile，然后将profile应用到对应的vEther interface上，如果profile修改，会立即对所有调用了该profile的vEther 口产生影响。

在N1K的VSM中定义port-profile，再将port-profile应用到N1K的vEth上。port-profile还分为up-link 和非up-link，连虚拟机的口应用非up-link的profile，连上游物理交换机的口应用UPLINK profile，uplink一般都是trunk。up-link profile还分为system port-profile(用于传递系统信息，比如VSM和VEM间的控制数据)和uplink port-profile(用于传递网络数据)然后将这些port-profile推送到vCenter，在vCenter中就会看到N1K中定义的port-profile，其表现为port group (相当于是一个端口池)，管理员在创建虚拟机时可以选择所需的port group应用到虚拟机的Vnic上，创建好的虚拟机启动后vNIC就会连接到这个port group中的一个接口。

2) 数据转发过程

同一母机下的数据转发由VEM本地完成，不出母机网卡。不同母机间通讯。。。。。

其所有的通讯过程就会一台物理的交换机完全一样。所不一样的就是他可以实现虚拟机的vMotion。

3) NEXUS 1000V下的vMotion

Nexus 1000V和VMware的vCenter是高度集成的，他能够完全了解所有服务器虚拟化活动，例如 VMware VMotion。

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

vCenter kick off 一个vMotion(可以是手动的也可以是自动处理的), 并发现一个通知信息 NEXUS 1000V交换机; vCenter转移VM文件到新的母机, N1K复制VM的端口状态信息(包括port policy等)到新的母机; VM文件移动完毕, 接口up,VM的vNIC和N1K port group 池中的一个接口发生Connect。网络接通后VM会通过ARP包宣告自己的MAC在网络中的新的位置。

这里有个问题, 在vMotion后虚拟机怎么知道自己和哪个vEth对应(重启也存在这个问题), 我想可能是在vCenter中用VM的UUID和profile作了绑定。只要vMotion后UUID不变, 其对应的profile也就不会变, 那么接到哪个(或什么样的)vEth也就确定了。

二、通过NIV实现VN-LINK

NIV通过硬件交换机(带VN-LINK功能)配合特殊网卡(带interface virtualizer功能, 比如思科的polo卡)来实现。带VN-LINK功能的交换机和带interface virtualizer功能网卡都支持VN-Tag技术。实际上就靠VN-Tag建立起了一条VN-LINK(其实就是标签技术)。

2. NIV 简单原理

NIV(network interface virtualization), 是CISCO提出的近似于硬件的解决方案, 是CISCO和VMware共同开发的虚拟机器网络协议, 可让Cisco VN-link功能在硬件中执行。他将所有的交换功能从母机(hypervisor)移走, 放到硬件交换机中, 从而与物理服务器无关。NIV模式的VIS(virtual interface switch)是不遵从802.1D的, 802.1D规定从某个接口收到的数据不可以再从该接口发出去, VIS必须允许此种数据转发才能实现同一母机下不同虚拟机之间的通讯。但是这种转发还是需要适当的标准来防止环路的发生的, 思科定义了VNTag协议用于VIS中, 利用Tag可以识别数据的来源和要去的目的是不是同一个地方。此协议标准已提交给IEEE802.3工作组。

NIV还需要物理服务器上有interface virtualizer功能来配合, 此功能可以通过hypervisor内建的软件实现, 也可通过带有此功能的网卡实现。interface virtualizer的功能就是在网卡上建立很多的vEth端口用来和VM的vNIC对应, 当vNIC有数据进入vEth时, vEth就会给他打上VN-Tag再发出去, 配合VN-Tag功能完成VN-link功能。也就是说NIV解决方案中需要用到interface virtualizer+VNTag+VN-link功能的交换机, 最终实现VN-link技术。

3.其工作流程如下:

1)VN-LINK形成

IV其实可以把他当着一个小的交换机来看待, 他具有很多的vEth口, 具备数据交换功能。在IV和VIS之间运CIV协议(Virtual Interface Control Protocol)。当IV中有新的vNIC建立时, IV使用这个协议请求VIS建立新的VIF, 如果要建的VIF在VIS中已经存在且符合IV的要求则不用再新建, 如果不存在就按照IV的要求新建一个(IV也可以请求VIS删除已建立的VIF)。新建完VIF后IV可以进一步请求VIS为某个vNIC enable对应的VIF接口, VIS enable完后会返回一个VIF ID给IV, IV将VIF ID和vNIC对应到自己的转发表中。这样vNIC和VIF就有了绑定关系, 就形成了一条VN-LINK。Link形成后还需要进行地址注册才可以使用该VIF, 也就是将vNIC的MAC关联到VIF(就好比是传到交换机的地址学习), 这样VIS才会跟据数据帧中目的MAC找到对应的VIF转发出去(并且会打上VN-TAG)。

猜测: 在IV的转发表中应该没有MAC的, 只需要根据tag做转发就可以了。

2) 数据转发过程

当有数据从服务器到网络时, interface virtualizer会识别数据的源vNIC, 并将从此vNIC产生的数据打上一个唯一的tag(即VN-tag)发送出去(VN-Tag封装源目的MAC之内, 此Tag的作用应该就是可以让VIS可以据此Tag来决定由哪个VIF来接收此帧或者也有可能是为了防止环路)àVIS收到数据后解出VN-Tag, 并根据转发表选择对应的出接口和出VIF, 并加上

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

新的VN-Tag（此Tag的作用应该是让目的端的interface virtualizer可以据此选择某个vEth作为出口，从而对应到相应的vNIC上）数据到interface virtualizer，interface virtualizer解出VN-Tag，移除VNtag 并将此数据分发到对应的Vnic。

interface virtualizer从来不会执行本地不同虚拟机之间的数据交换，交换处理被从hypervisor中解放出来放到物理网络设备上执行。和interface virtualizer相连的交换机称为virtual interface switch (VIS)，此交换机不仅负责不同物理接口间的数据交换，还负责不同virtual interfaces (VIFs)间的数据交换，VIF和vNIC间会有对应关系(通过vEth)，不同VIF间的数据交换其实就是不同虚拟机间的数据交互。任何数据交换和策略的实施都是在VIS中，并不是在hypervisor。VIS可以是任意的支持NIV(network interface virtualization)功能的交换机。

4. NIV中的VMotion
 - 1) vMotion 目前只能在同一个二层网内发生
 - 2) vCenter发送vMontion包（从源母机发向目标母机）
 - 3) 虚拟机的状态信息被copy到目标母机
 - 4) 虚拟接口(vEth)状态信息被相应copy到目标母机上
 - 5) 通过RARP广播完成MAC在网络中的迁移（从一个理交换机端口到另一个交换机）

第八节 N1000---VM ESX Server

VMware ESX 服务器是在通用环境下分区和整合系统的虚拟主机软件。它是具有高级资源管理功能高效，灵活的虚拟主机平台。

VMware ESX Server 为适用于任何系统环境的企业级的虚拟计算机软件。大型机级别的架构提供了空前的性能和操作控制。它能提供完全动态的资源可测量控制，适合各种要求严格的应用程序的需要，同时可以实现服务器部署整合，为企业未来成长所需扩展空间。亦提供储存虚拟化的能力。除可因兼并服务器减少设备购买及维护成本外，亦可因效能的尖峰离峰需求，以VMotion技术在各服务器或刀片服务器之刀板间弹性动态迁移系统平台，让IT人员做更有效的资源调度，并获得更好且安全周密的防护，当系统发生灾难时，可以在最短的时间(毋需重新安装 操作系统)，迅速复原系统的运作。

VMware ESX服务器传输物理系统到逻辑资源池，系统和应用程序被分离在包含多虚拟机的单一硬件片上。系统资源按需动态分配到任何操作系统上，提供大型机的性能和服务器资源控制。VMware ESX 服务器通过在安全便捷的虚拟机上执行分区和分割服务器资源来简化服务器的基础构架。VMware ESX 服务器确保服务器资源能被远程管理，自动预制和标准化。高级资源管理控制通过企业担保服务质量。

1.ESX Server 2.5.x特性概述

ESX Server 2.5.x有下面的特性

ESX server物理机

- 每ESX Server支持16个逻辑CPU
- 每ESX Server支持80个虚拟机CPU
- 每ESX Server支持64GB内存
- 支持8个swap文件，每个swap文件的最大容量为64GB
- 每系统支持64个适配器

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

- 支持8个千兆网卡或者16个百兆网卡
- 每个虚拟交换机支持32个虚拟机
- 每ESX Server支持16个光纤通道卡
- 每个存储阵列支持128个LUN
- 每ESX Server支持128个LUN

ESX Server 2.5.x虚拟机

- 通过可选的vSMP模块，每个虚拟机可以支持2个虚拟CPU
- 每个虚拟机支持3.6GB内存
- 支持4个虚拟scsi适配器并且可以支持到15个scsi硬盘
- 虚拟磁盘容量支持9GB
- 支持4个虚拟网卡

关于最新的支持操作系统和硬件的列表可以参考下面链接提供的系统兼容向导。

http://www.vmware.com/vmtn/resources/esx_resources.html

2.ESX Server 2.5.x的虚拟结构

ESX Server 2.5.x的虚拟结构由下面几个部分组成

- ESX Server 2.5.x
- VirtualCenter 1.3
- vSMP
- Vmotion

ESXServer运行在一个物理机上。VirtualCenter即可以运行在物理机上也可以运行在一个虚拟机上。如果你决定将VirtualCenter运行在一个虚拟机上，你需要考虑一个问题，如果虚拟机所属的ESXServer离线了，你将不能访问VirtualCenter直到ESXServer重新上线或者你将虚拟机重新运行在另一个机器上。vSMP和VMotion是已经安装的特性但是需要一个license去激活。

VMware提供了一个VIN (Virtual Infrastructure Node) license，内容如下

- ESX Server license
- Virtual SMP license
- VirtualCenter Agent license
- vMotion license

购买VIN license比单独购买各个license更节省成本。

3.服务器的规模

使用服务器的规模由以下几个因素决定

- 目前方案的规模
- 将来发展的估计
- 高可用性和容灾计划
- 预算的限制

有一些不同的方法去估算你将需要的ESX服务器的数量，下面介绍最流行的两种：

- 最容易的方法是计算每个4-5个虚拟CPU对应一个物理CPU

这个结果是大概一个4路服务器对应16 – 20个虚拟CPU，一个8路服务器对应32 – 40个虚拟CPU，这里假设每个虚拟CPU运行的负载量适中。

关于内存方面，如果你给每个虚拟机分配1GB内存，在大多数情况下已经足够处理虚拟机运行，管理终端和虚拟管理了。如果你计划在虚拟机上运行一些内存敏感的应用，应该适当的加大这个值。

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

根据上面的估算，我们可以用一个8路（2节点）、配置32GB内存的x3950支持32个虚拟机，用一个16路（4节点）、配置64GB的x3950支持64个虚拟机。

上面的估算都是基于一个单核CPU，因为一个双核CPU不能提供100%的相当于两个单核cpu运行性能，我们建议你按1.5个单核CPU来估算1个双核CPU，也就是每双核CPU对应6-7个虚拟机。

· 如果你有一些运行应用的物理机，也可以通过计算CPU利用率平均峰值来换算成MHz总数来进行虚拟机数量的估算。例如，如果你有两个500MHz的CPU，CPU利用率平均峰值为50%，那么你的虚拟机CPU计算结果应为500MHz。

为了得到CPU利用率的平均峰值，你需要对CPU的利用率在运行着应用的情况下至少进行一周的记录，为了计算结果更准确建议进行一个月的记录。如果你已经有了企业级的监控软件比如IBM Tivoli、HP OpenView、NetIQ等等，那么你已经可以直接通过这些软件得到你想要的信息。

下一步你需要把你所有运行ESX Server的服务器的主频加起来。例如，一个配置3GHzCPU的2节点8路的x3950服务器计算结果为24000MHz。计算结果先需要减去10%，这部分是终端OS需要的，结果是21600MHz。

因为附加的占用率和过载的情况下需要减去一些数量，通常20%是安全的数值。得到了17280MHz可以提供给虚拟机的应用运行。分为每个虚拟机500MHz，结果为大概可以运行34个虚拟机(17,280/500=34.5)。

你可以用类似的方法来估算虚拟系统所需要的内存总量。首先在物理机上得到平均内存利用率，然后每个虚拟机加上54MB为了虚拟化的开销，为内存平均利用率超过512MB的系统加上32MB内存，这些就是你的虚拟机需要的内存总量。还需要加上分配给Service Console的内存（一个8路的ESX Server系统这个值一般设为512MB），加上VMkernel所需的24MB内存，这就是整个系统需要的内存总量。

例如，你有10个平均内存利用率为512MB的物理机需要转换为虚拟机总量为5120MB，加上每个虚拟机54MB的虚拟化开销（5120+540=5660 MB），这是虚拟机需要的总量。再加上Service Console需要的内存（5660+512= 6172 MB）和VMkernel的24MB内存（6172+24=6196），得到最后的结果是这10个虚拟机需要6GB内存。

我们看到上面两种方法在8路的x3950服务器上得出的结果很相似，其实按照一般的经验通常上面两种方法估算的结果都是很接近的，所以为了节省时间，我们建议使用第一种方法来初步规划你的ESX Server的规模。

未来方案规模的发展是很难去决定的。经常出现这样的情况，当一些公司第一次执行VMware的虚拟环境的时候先部署了一部分虚拟机，但是使用后发现了虚拟环境带来了许多的好处并且没有降低系统的运行效率，这样新的虚拟机的数量迅速增加，超过了现有的CPU、内存和I/O资源，这将导致总体的运行效率收到影响。所以需要为虚拟环境的扩展提前作出一定的考虑。

4.关于VMotion

在设计虚拟环境的时候，VMotion需要被做为一个重点去考虑。VMotion特性可以实现在虚拟机保持运行的情况下实现虚拟机在物理机之间的转移。因为VMotion需要在两个物理机之间转移系统运行的一些底层的状态，所有要求两个物理平台的CPU必须运行一样的指令集。这就意味着执行VMotion操作的物理服务器的CPU必须符合下列的要求

- 同厂商的CPU产品(Intel or AMD)
- 同系列的CPU产品(Pentium III, Pentium 4, Opteron, etc)

有时同样系列的CPU产品在扩展特性上会有很大的不同，比如64位扩展技术和SSE3。在这

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

种情况下不能实现VMotion的操作即使是同样系列的CPU。CPU的主频和缓存不是考虑的主要因素，但是一些扩展特性的不同可能导致在物理机之间的VMotion操作失败。

例如，x366和x260服务器和x3950使用同样类型的CPU，它们适合与x3950配置VMotion操作。然而其他使用不同CPU的xSeries服务器就不能了。

VMotion要求准备一个专署的网卡，VMware要求提供一个千兆网卡供VMotion使用。但是如果你只有百兆网卡可用，也可以配置VMware使用百兆网卡，当然这样VMotion的时间将显著增加。另外一个重要的要求是VMotion要求一个共享的存储。

5.计划你的Server farm

在VirtualCenter 1.x中，一个farm就是一个组成你的虚拟环境的ESX Server系统的组。一个farm也是一个VMotion的划分，这意味着相互配置了VMotion操作的服务器必须在同一个farm中。在计划的过程中，你需要考虑每个farm中主机的数量。下面是VMware帮助你划分farm的一些参考：

- VMware建议连接到一个单独VMFS卷上的ESX Server不要超过16个。
- VMware建议每个LUN不要划分超过32个I/O敏感的虚拟机，或者100个对I/O要求很低的虚拟机。
- 每个VMFS卷不要超过255个文件。
- 存储容量有2TB的限制。

因为VMotion要求共享的存储，那么将限制每个farm可以有16个ESX Server。你可能希望去创建一个小一点的farm，那么可以创建一个使用VMotion的一个只有两个ESX Server的farm。

6. 存储容量

就像服务器规模的估算一样，没有一个普遍的答案应用到每个不同的环境中。前面部分已经提到了每个VMFS卷运行I/O需求敏感的虚拟机的数量不要超过32个，应该尽量的减少资源争用和SCSI锁情况的出现。

有很多方法去决定VMFS卷的合适大小，下面是最简单的方法。

如果你决定用2个8路的x3950服务器，每个服务器运行32个虚拟机将符合对CPU性能的要求。用前面提到的32个虚拟机每LUN的策略来看，你需要为这些虚拟机建立两个LUN。如果你建立新的虚拟机可以先估算一下虚拟硬盘的平均容量。如果每个虚拟机平均需要20GB硬盘，这样每个LUN需要640GB。你还需要加上一些附加的容量为了以后的扩展需要，通常10%是个合适的比率，加上这部分结果为720GB。如果你计划使用redo日志，需要相应的多划分一些空间。

7. 规划网络

在服务器farm设计网络部分的时候有很多选项，对这些选项的选择一般都是基于你的物理网络或者公司的网络安全策略。重要的一点是ESX Server强烈建议使用千兆网络连接。

在ESX Server 2.5.x配置网络的时候，你需要考虑三个基本的组成部分。

- Service console

出于对运行效率和网络安全的考虑，VMware建议你为Service console提供一个专用的网卡。你可以把service console网卡连接在你的数据中心的专用的管理网络中。

在默认的配置中，一个百兆网卡足以提供service console所需的带宽。如果你计划用service console运行备份或者其他需要高带宽的应用，就建议使用千兆网卡。

- 虚拟机

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

虚拟机使用一个与service console分离的网络。虽然系统并没有强制要求使用千兆网卡，但在实际应用中我们建议你使用千兆网卡提供虚拟机的网络连接，因为32个虚拟机将产生巨大的网络通讯量。

一个比较好的计划是10-20个虚拟机使用一个千兆网卡。这意味着一个8路x3950运行32个虚拟机最少需要2个千兆网卡。请注意这是最小的建议，再添加1个或2个千兆网卡将为虚拟机提供更充足的带宽。

另外一个需要考虑的是VLAN的配置，假设你的数据中心里有多个VLAN需要应用到虚拟机上，这种情况下你有两种不同的选择：

第一种是为每个物理的VLAN在ESX Server上配置一块网卡，当你的网络环境比较简单的时候可以使用这个选择，但是如果你有10个网络这个选择就变的不可行了。注意ESX Server 2.5.x仅仅支持最大8个千兆网卡。

另一种是使用ESX Server对VLAN (802.1q) 的支持，这个选项为每个VLAN创建一个虚拟的端口组在虚拟交换机上。如果你的物理交换机支持，建议你使用这种方法。

在配置中还需要考虑虚拟机网络的冗余连接，ESX Server 2.5.x不仅支持多个网卡连接到一个虚拟交换机提高带宽，也提供了对网卡失败的冗余切换。为了实现高可用性我们建议你对网卡进行绑定。

· VMotion

VMware要求为VMotion提供一个独立的千兆网卡。不推荐使用百兆网卡，因为这将大大影响运行效率。你应该为VMotion单独分配一个千兆网卡并且划分单独的子网。

如果你只有两个运行ESX Server的服务器，你可以使用交叉线直接连接两个服务器的网卡作为虚拟网络使用，这个方法也可以在排除VMotion的故障的时候使用。

8.网络负载均衡

ESX Server 2.5.x为虚拟机提供了两种网络负载均衡的方法。

· 基于MAC地址的网络负载均衡是ESX Server默认的方式。这种方法简单的将两个物理网卡连接到虚拟交换机不需要在物理交换机上进行额外的配置，缺点是负载均衡的效率不是非常理想，这种方法有时使一个虚拟机经常使用某一个物理的网卡，并且不能手动的调整每个虚拟机使用的物理网卡。

· 基于ip的网络负载均衡，这种方法提供了一种网络负载均衡的更好选择，缺点是你需要在你的物理交换机上进行额外的配置，需要配置你的物理交换机的802.3ad协议（或者Cisco交换机的EtherChannel）。这是一个实现高运行效率的选择。

以上是对ESX Server 2.5.x的网络的概述。一些高级的应用，比如备份网络、DMZ网络和配置的细节超出了这本红皮书的介绍范围。如果需要更深入的信息请访问VMware的官方网站，链接如下

http://www.vmware.com/support/pubs/esx_pubs.html

http://www.vmware.com/vmtn/resources/esx_resources.html

第九节 许可证

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

整体许可证:

Cisco NX-OS有三个许可证级别。

1: 基本许可证提供丰富的特性集, 它与硬件捆绑在一起, 免费提供。企业许可证支持适用于许多企业部署的高级功能, 而高级许可证则支持下一代功能, 如VDC和Cisco TrustSec。

2: 企业软件包

下列功能仅通过Cisco NX-OS 4.0的企业许可证提供:

- IP路由
- OSPF v2和v3 (IPv4和v6)
- IS-IS
- BGP
- EIGRP
- IP组播
- PIM: 稀疏模式(IPv4和IPv6)、双向模式和SSM模式
- MSDP
- PBR
- GRE隧道

3: 高级软件包

高级许可证支持Cisco NX-OS 4.0的下列功能:

- VDC (OTV必须VDC支持)
- Cisco TrustSec

订购信息

产品名称 产品编号

Cisco NX-OS 企业局域网许可证 N7K-LAN1K9

Cisco NX-OS 高级局域网许可证 N7K-ADV1K9

Cisco NX-OS 4.0软件, 用于Cisco Nexus 7000 Supervisor 1 N7KS1K9-401A1.1

板卡许可证:

M和F系列板卡中: non-XL和XL许可证的区别是 无XL许可在最大MAC地址数目、最大路由条目等等都均有限制, 比如non-XL只能最大128K路由, XL可以更多。

注意, 是一张板卡, 无论多少VDC在其中。

图形化管理许可证:

DCNM, 默认基本许可, 购买高级许可后, 方可管理所有特性 (VDC FP等等)。

N5K许可证:

N5010-SSK9(=)

Nexus 5010 Storage Protocol Services License

N5020-SSK9(=)

Nexus 5020 Storage Protocol Services License

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

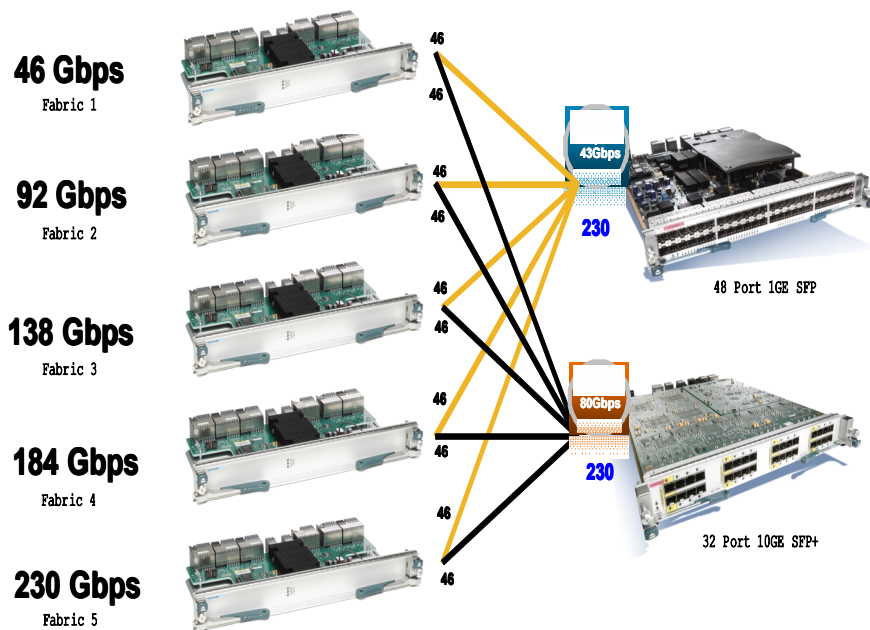
victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

N5000FMS1K9(=)

Nexus 5000 Fabric Manager Server License

第十节 交换矩阵

用于光纤连接的2个I/O模块



一代：目前CISCO建议如果要实现官方标称的最大背板带宽以及每插槽最大带宽，配置三个光纤矩阵模块，以提供N+1的备援

二代：最大可提供550Gb/s

第十一节 M 和 F 系列板卡

M-module提供L3(本模块内交换) F-module提供L2 (提升到SUP)

M不提供FCOE,F提供

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

N7K-F132XP-15 每两个口就有自己独立的转发模块 (SOC)，特别注重L2交换性能，如果F模块要实现L3功能，只能送到M模块帮助进行 (如果没有M模块，只有SUP1呢？回答是不行，一定要有M模块，SUP1不提供L3，所以要F模块在，实现L3的话，必须一起配置M模块，没有M模块，连inter vlan都不能建立)。PS：尽量将同一个VLAN放到F模块上。

问：那是否整个N7K，没有配置任何一块M模块，只有F模块以及N5000，是否就无法实现L3的转发？现在思科的回答是：是的要L3转发，必须配置M模块)

non-XL和XL许可证的区别是 无XL许可只能最大128K路由，XL可以更多。意思就是如果一张 不带XL许可的板卡上有3个VDC，之和是超过128K路由，是不可以的。但是可以通过去掉一个VDC,一张板卡放2个VDC，不超过128K来调剂。

第十二节 电源

电源：

fu red 不考虑冗余，全功率可以达到 18KW，每一个电源提供6KW (2路，每路3KW) 12KW ，可以N+1,1个电源做备份。6+6=12 6做备份

grid redundancy 9KW 两路电源组接入， 每个电源1路主，一路备份，所以3+3+3=9KW 这种模式每路有自己的电源组

默认fu red 9KW 一路线接入，每个电源1路主，一路备份

用于 *Nexus 7010* 的 6Kw 电源

全冗余 可用功率 = ~9Kw 备用功率 = ~9Kw	(6x 220v = 18Kw) Can Lose 1 PS or 1 GRID	
GRID 冗余 可用功率 = ~9Kw 备用功率 = ~9Kw	(6x 220v = 18Kw) Can Lose 1 GRID	
N+1 冗余 可用功率 = ~12Kw 备用功率 = ~6Kw	(6x 220v = 18Kw) Can Lose 1 PS 默认状态	
组合型 可用功率 = ~18Kw 备用功率 = 0Kw	(6x 220v = 18Kw) 没????!	

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

电源需求:

7010 需要2400w 启动

7018 需要2193w 启动

组件	最大功率	一般功率
10插槽风扇架	1680w	300w
18插槽风扇架	1273w	569w
10槽光纤模块	60w	55w
18槽光纤模块	100w	90w
管理引擎	210w	190w
32口万兆以太网模块	750w	611w
48口10/100/1000M模块	400w	358w
48口SFP模块	400w	358w

第十三节 VDC

虚拟设备，将一台划分为几台。(一台N7K 最多4个VDC)

VD1 为官方默认虚拟交换机，并且官方不太建议使用VD1来运行业务数据。

VDC1是预设的，永远激活的，不可删除的，默认必有的。在VDC1里面建立其余3个VDC。

你第一次登录一台初始化的N7K，就等于进入VDC1，VDC1是管理VDC。

所有VDC的 VLAN和路由表均完全独立，包括任何故障都是独立的。(除非整个N7K-OS平台死掉)

在VDC1中执行reload命令会影响所有VDC.其余的VDC reload不会影响其他VDC。

某些资源又可以在VDC们中可以共用，比如boot image\SW名称\NTP\COPP\IN-BAN SPAN SESSIONS等

意思是可以共用资源，不是说只能统一设置，比如你可以不同的VDC使用不同的IMAGE启动。

物理接口都是专用的，但是除外的是OOB以太网管理通信端口是共享的。

可以建立资源模板，供不同的VDC调用。已经在VDC中使能模板后，再改动此模板将无作用，需要再使能后才可起作用。

只能在VDC1里使用 switch VDC 命令转换到其他VDC，其他VDC不能再切了，只能先切回VDC1，再切其他VDC

两块主引擎和一块主引擎在管理VDC上有一定的区别，当出现VDC出现故障时：1张的话，只能restart，系统自动删除VDC并重新建立VDC。或者reload,重新加载管理模块。2张的话，可switchover，启动管理模块的switchover.也可restart。思科官方建议单一VDC restart,不建议switchover，因为会影响所有的VDC。而且VDC1尽量只做管理。

VDC1里面如果有改动资源分配，必须保存一下，其余的VDC才能执行本身的配置保存。

VDC试用许可是120天，过期全部的VDC配置将会被删掉。

VDC1是预设的，管理VDC，所有创建、删除的VDC的操作都在VDC1中进行。

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

VDC1可以将接口移动到VDC2，但是不能再将这个接口从VDC2移动到VDC1，只能直接指定这个接口重新分配到VDC1。

第十四节 VRF

NX-OS中的VRF，不是用在MPLS VPN中的VRF。

NX-OS中，一个VDC默认自带2个VRF，一个是用于数据转发的VRF，一个用于管理层面的VRF。

第一个用于数据的VRF默认不带VRF的参数，而用于管理层面的VRF就带VRF参数

也就是说，进入一个VDC，就开始配置路由协议，是默认属于第一个数据VRF的。

而mgmt0口就属于管理VRF，并且能够被任何一个VDC调用。

每一个VDC都可以起多个VRF，而且理论上VRF可以同名，前提是属于不同的VDC。

建立VRF命令：vrf context XX

分配命令：vrf member (建立VRF后，再使用此命令将L3，L2等接口资源分配给该VRF)

举例：

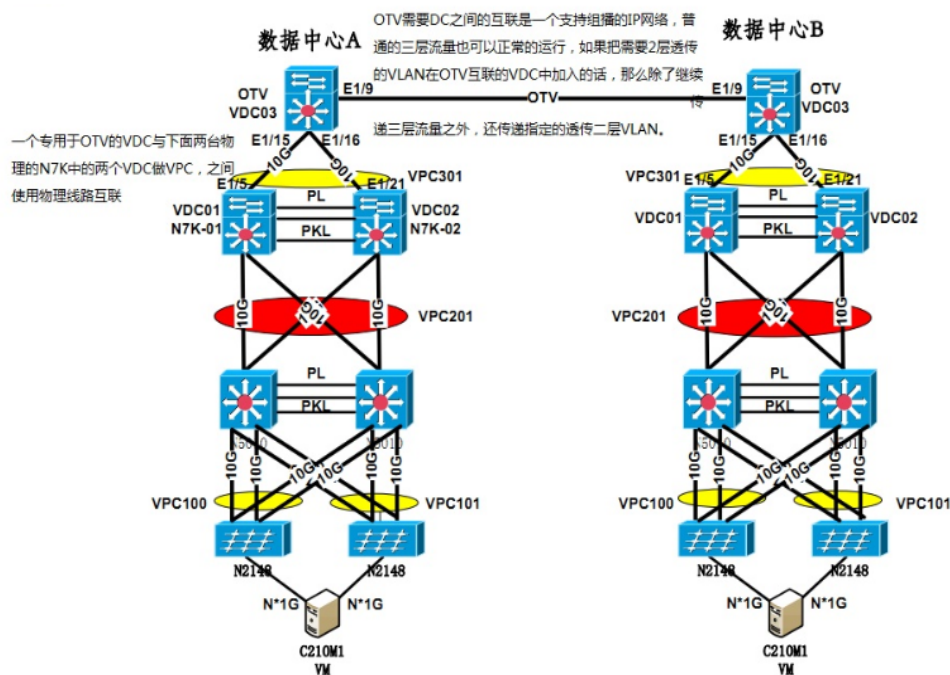
```
N7K-1(config)#vrf context TEST
```

```
N7K-1(config)#interface vlan 10
```

```
N7K-1(config-if)#vrf member TEST
```

```
N7K-1(config-if)#ip add 10.10.10.1/24
```

第十五节 实验举例



本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

Step 1. 初始化 VDC03

建立VDC03，为 VDC03 分配物理接口资源，配置 VDC03 管理接口地址及管理口令 DCcisco123，配置建立 VLAN 101-104、开启LACP 特性。

DC-A-N7010中建立 VDC03

```
vdc VDC03
  allocate interface Ethernet1/9-16
  allocate interface Ethernet3/13-24
```

配置 VDC03 管理接口地址及 VRF

```
vrf context management
  ip route 0.0.0.0/0 10.75.58.1
interface mgmt0
  ip address 10.75.58.172/24
```

配置建立 VLAN 101-104

```
vlan 101-104
```

配置开启 OTV及 LACP 特性

```
feature otv
feature lacp
```

Step 2. 配置 VDC03 与 VDC02/01互联：

配置 E1/15-16 为Port-channel 20接口下联 VDC01和 VDC02

```
interface port-channel20
  switchport
  switchport mode trunk
interface Ethernet1/15-16
  switchport
  switchport mode trunk
  channel-group 20 mode active
  no shutdown
```

分别配置VDC01 E1/5 及VDC02 E1/21 隶属于 Port-channel 301 及VPC301。

VPD01配置如下：

```
interface port-channel301
  switchport
  switchport mode trunk
  vpc 301
interface Ethernet1/5
```

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

```
switchport
switchport mode trunk
channel-group 301 mode active
no shutdown
```

VPD02配置如下:

```
interface port-channel301
  switchport
  switchport mode trunk
  vpc 301
interface Ethernet1/21
  switchport
  switchport mode trunk
  channel-group 301 mode active
  no shutdown
```

Step 3. 配置启用 OTV

配置 VDC03 背靠背接口 E1/9 IP 地址并且启用组播
数据中心 A 的 VDC03 E1/9: 10.1.1.1/24.
数据中心B 的VDC03 E1/9: 10.1.1.2/24.

数据中心A VDC03 配置如下:

```
interface Ethernet1/9
  ip address 10.1.1.1/24
  ip igmp version 3
  no shutdown

interface Overlay1
  description site A
  otv join-interface Ethernet1/9
  otv control-group 239.1.1.1
  otv data-group 232.1.1.0/28
  otv extend-vlan 101-104
  no shutdown
```

数据中心B VDC03配置如下:

```
interface Ethernet1/9
  ip address 10.1.1.2/24
  ip igmp version 3
  no shutdown
```

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

```
interface Overlay1
  description site B
  otn join-interface Ethernet1/9
  otn control-group 239.1.1.1
  otn data-group 232.1.1.0/28
  otn extend-vlan 101-104
  no shutdown
```

第十六节 交换

NX-OS系列不支持DTP，必须手动配置指定的静态trunk，TRUNK仍然默认允许所有VLAN通过。两边对联设备本真VLAN必须一致。

N7K中，所有VDC VLAN之和不能超过16000个。也就是4个4000

NX-OS 根据对端STP切换到原老的STP协议。实际应用中，却有一些问题。但是如果对端老设备也配置支持RSTP或者PVST，就没有问题。

整个交换网络，要么整体MSTP，要么整体PVST，不能夜航过渡。

MSTP中，VLAN对应表一定要统一，名字也要一致。

span port type edge 代替 span portfast

bridge assurance代替 loop guard来防止交换环路。使用BPDU在每个接口上发送，当作hello，可以发现潜在回路。所以每一个设备的环路互联接口也都必须配置bridge assurance

也就是说，在每一台设备互联接口下：spanning-tree port type network是必配的

第十七节 VLAN

vlan1:是预设VLAN不能被修改和删除

VLAN2-1005: 属于能够被建立、使用、修改与删除的一般性VLAN

VLAN 1006-4094: 属于能够被建立、使用、修改与删除的延伸性VLAN，这些VLAN的状态永远是使用中，而这类VLAN会一直启用并且不能被停用。

VLAN 3948-4047和4094: 仅作为内部使用

由于VDC的存在，VLAN 编号能够在不同的VDC上重复使用，但是所有VDC VLAN数目之和不能超过16000个。

```
N7K-1(config)#vlan 20
```

```
N7K-1(config-vlan)#exit
```

```
N7K-1(config)#switchto vdc RED
```

```
N7K-1-RED#config
```

```
N7K-1-RED(config)#vlan 20
```

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

N7K-1(config-vlan)#

第十八节 端口聚合

端口聚合将多个以太网连接聚合成一个单独的逻辑以太网连接。

端口聚合常用于提高可用性和带宽。

每台设备所有的VDC支持最多256条端口聚合。

NX-OS中，端口聚合只支持两种方式：

1: 静态模式(无协议)

2: 链接访问控制协议(LACP) - 802.3ad

NX-OS现在只支持LACP，不支持PAGP。而且只能静态手动配置。

LACP中，配置为ACTIVE模式才可以协商，passive模式不会协调。通信端口ON模式为静态建立，不会有任何协调。(需要LACP协调的才可实现 线路故障冗余)

NX-OS中：P-C的负载均衡选择大幅增加：

n7000(config)# port-channel load-balance ethernet ?

dest-ip-port	Destination IP address and L4 port
dest-ip-port-vlan	Destination IP address, L4 port and VLAN
destination-ip-vlan	Destination IP address and VLAN
destination-mac	Destination MAC address
destination-port	Destination L4 port
source-dest-ip-port	Source & Destination IP address and L4 port
source-dest-ip-port-vlan	Source & Destination IP address, L4 port and VLAN
source-dest-ip-vlan	Source & Destination IP address and VLAN
(Default for IP)	
source-dest-mac	Source & Destination MAC address (Default for Non-IP)
source-dest-port	Source & Destination L4 port
source-ip-port	Source IP address and L4 port
source-ip-port-vlan	Source IP address, L4 port and VLAN
source-ip-vlan	Source IP address and VLAN
source-mac	Source MAC address
source-port	Source L4 port

L2的P-C配置举例：

n7000(config)#feature lacp

n7000(config)# interface ethernet 1/25,ethernet 1/27

n7000(config-if-range)#switchport

n7000(config-if-range)#channel-group 1 mode active

n7000(config)#interface port-channel 1

n7000(config-if)#switchport mode trunk

n7000(config-if)#以及其他的L2命令

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

L3的P-C:

```
interface port-channel1
  ip address 192.168.10.1/24
interface Ethernet1/13
  channel-group 1 mode active
interface Ethernet1/14
  channel-group 1 mode active
查看P-C状态
show port-channel summary
```

第十九节 MSTP

NX-OS中，为了最大兼容性，CISCO建议全网交换设备一定要统一一种协议，不能类似普通以太网环境这样STP\RSTP\PVST这几种协议均有并兼容状态。

即使全CISCO交换设备，如果NX-OS使用STP，6509使用PVST也会有不稳定的情况。

所以CISCO推荐全网一致采用MSTP协议。

```
N7K-1 (config)#spanning-tree mod mst
N7K-1 (config)#spanning-tree mst configuration
N7K-1 (config-mst)#name MST
N7K-1 (config-mst)#revision 10
N7K-1 (config-mst)#instance 1 vlan 1,100
N7K-1 (config-mst)#instance 1 vlan 101,200
N7K-1 (config)#spanning-tree mst 1 root primary
N7K-1 (config)#spanning-tree mst 2 root secondary
N7K-2 (config)#spanning-tree mst 2 root primary
N7K-2 (config)#spanning-tree mst 1 root secondary
```

第二十节 UDLD

UDLD特性一旦被启用，所有10G口会自动执行UDLD，1G的需要手动开启。

```
N7K-1(config)#feature udld
N7K-1(config)#udld aggressive    \\全域性的启用积极模式。
```

第二十一节 VPC

跨设备的以太通道，改进STP，VPC最多只能2个上联设备。

VPC，下方一个设备可以是非N7K系列设备（C65，PC server 等），配置也只是普通以太通道即可。

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

VPC本身只是L2的port-channel,不支持L3的方式。

N7K系列最多可以16个捆绑(全N7K设备互联,如果和普通设备则为8个),并且N7K系列的以太通道可以支持源目IP以及TCP 端口来负载均衡等

但是这16个中,只有8个是启用的,其余8个是处于热备份状态。

16个中,不是论单对单,而是在2对2互联的情况下的一共16个,也就是说,一台N5和2台N7还是只能8个,要2台N5和2台N7互联才可以最多16个。因为2个N5的8个虚拟为一个16了

如果VPC的两个PEER其中一个当机后,端口会变成orphan port模式(孤立状态),也就是普通模式。

VPC后,只有主交换机发BPDU到线路。

VPC只能在不同的物理SW上做。不能在属于一台N7K中,2个VDC之间和对端交换机做。

VPC端口本身只能配置为sw port,L2模式。L3你可以使用inter vlan

VPC指定对端的IP建议使用管理IP。N5K和N7K都不建议两个设备之间用双绞线对联管理口,而是通过正常的L3模式连接。

VPC的主要构件:

peer-link

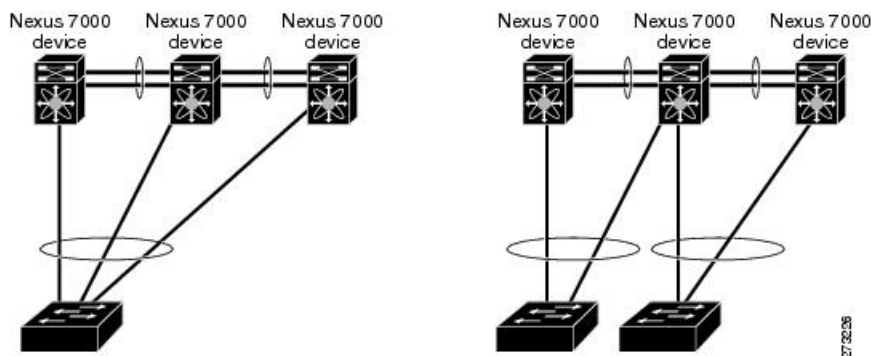
Peer keepalive link

VPC的优化组件:

peer-gateway

peer-switch

VPC不允许的构架:

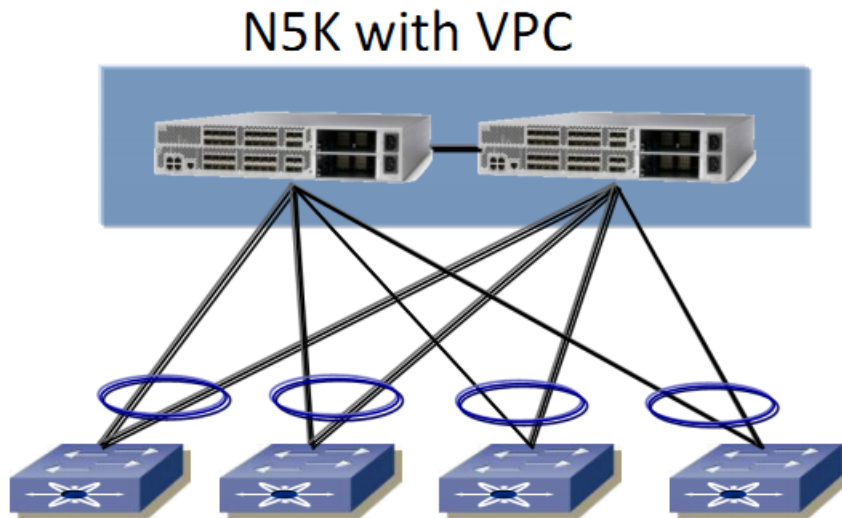


允许的架构:

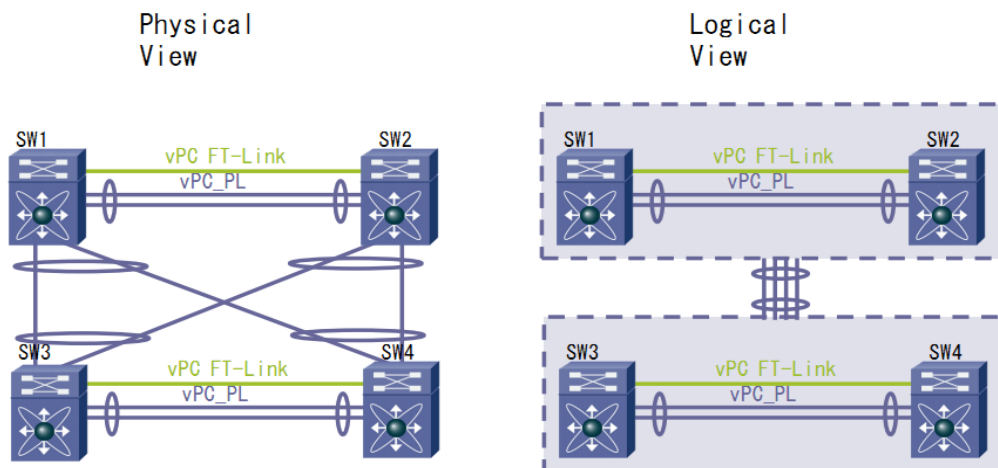
本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



Standard Switch with Port-Channel Support



第二十二节 Peer keepalive link

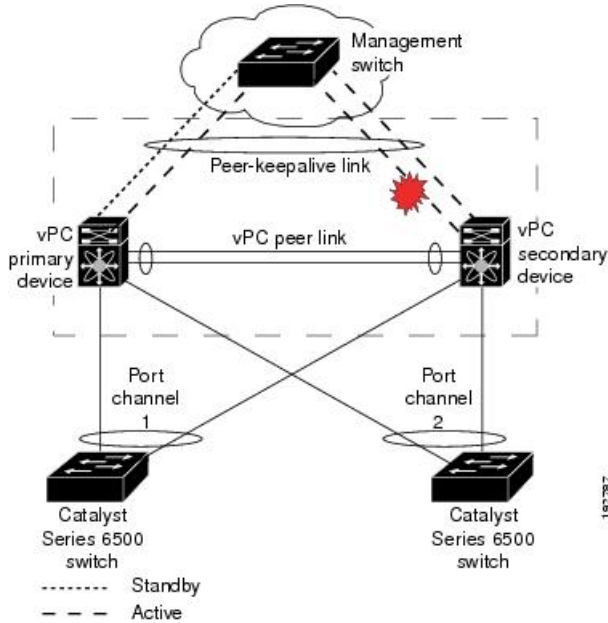
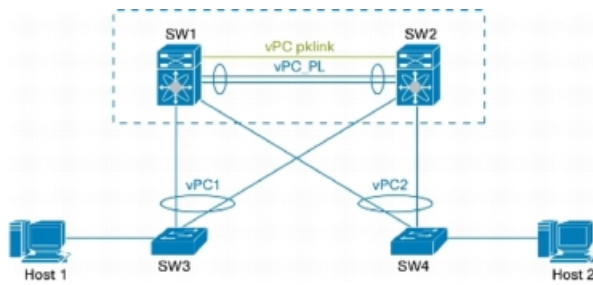
Peer keepalive link: 通过L3层面（OOB，或者说只要L3可达就行）检测对方是否存在，只是用于检测是否对方激活，状态同步由peer-link完成。并且不能合用，要分开物理线路。2者都要配置通过CFS协议。

目前思科建议使用专用的L3网络，比如管理口的专网，或者其他L3的互连网络，不建议使用双绞交叉线进行直连，否则会出现一定的问题。

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



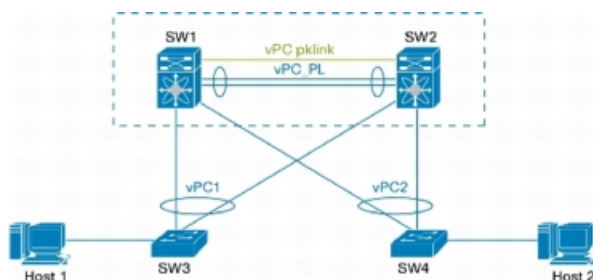
第二十三节 peer-link

peer-link:VPC在两台N7K上的统一，直接连到对方，使用 cfs协议（peer-link的物理线路是既可以传递同步状态，也可以传递客户业务）

PEER-LINK一定只能是10G口。

注意：运行PEER-LINK 的线路不能承载Peer keepalive link，两者必须分开。

实际中，建议Peer keepalive link使用专有的L3层网络（比如管理专网），而PEER-LINK就使用设备互联的以太通道线路即可。



本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

第二十四节 peer-gateway and switch

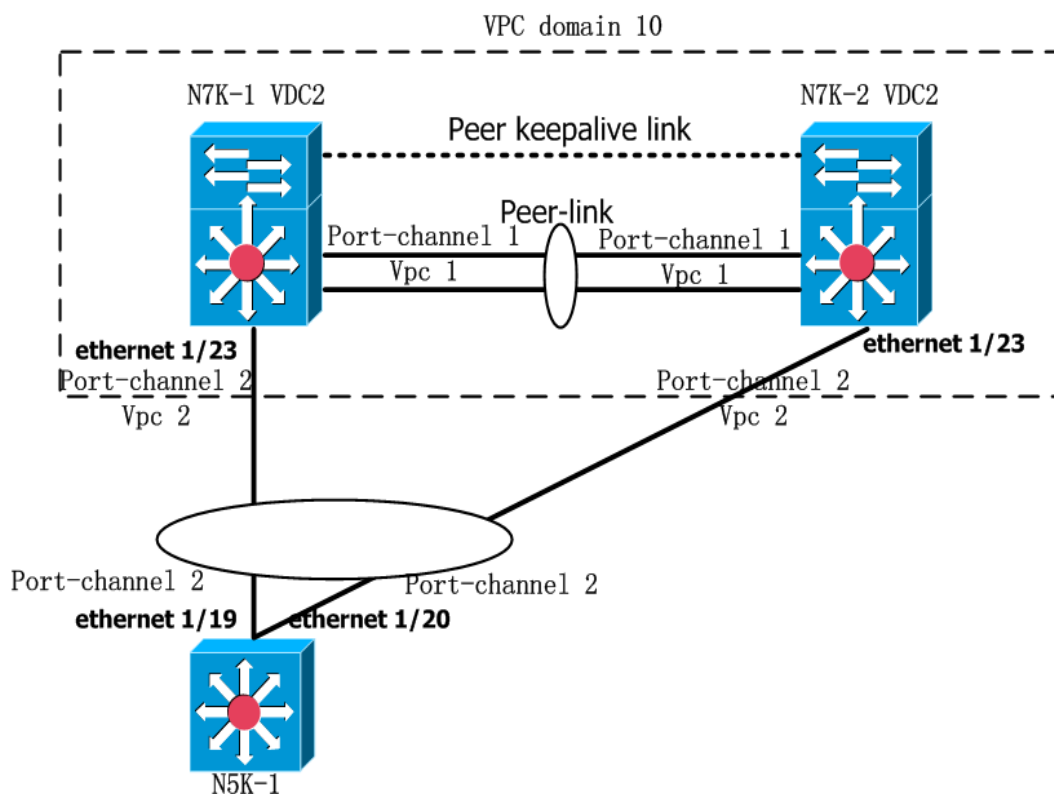
peer-gateway: 当上联一个设备, 做通道没有问题, 但是当上联2个交换设备时, 主机发数据包如果分别发到2个交换机是不恰当的(因为包发到备份交换机, 会转到主处理)。而使用peer-gateway来优化后, 包如果发到备交换机, 备SW可以直接处理。这种情况。有点类似HSRP的模式。

peer-g不是必配, 而是优化。

peer-switch 从L2的BPDU包上进行了优化, 启用之后, 两台设备都发BPDU, 并且BPDU中bridgid一致, 可以优化L2的链路收敛问题。

第二十五节 配置 VPC (实验)

实验TOP图:



注释: N5K-1可以为C6509之类的普通思科交换机

步骤: (本步骤前假定已经配置好生成树协议)

- 1: 配置两台N7K的2条互联线为普通的以太通道
- 2: 配置两台N7K与N5K的互联线为普通的以太通道
- 3: 两个N7K交换机设定统一的vpc domain ID (指定对等体, 组成一个虚拟VPC组)
- 4: 设定Peer keepalive link,并确认启用。

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

5: 设定peer-link,并确认启用。

6: 可以启用Peer-gateway 和Peer-switch 来优化 VPC

7: 针对不同的以太通道配置不同的VPC组号。(在针对N7K互联, 以及N7K-N5K互联的以太通道组里, 配置VPC组号, 两台N7K中, 用于互联的以太通道组中的VPC组号要一致, 两台N7K中, 分别连接至N5K的以太通道组的VPC组号要一致)

1: 先配置两个设备的2条上联线为普通的以太通道

```
N7K-1(config)#feature lacp
```

```
N7K-1(config)# interface ethernet 1/25,ethernet 1/27
```

```
N7K-1(config-if-range)#switchport
```

```
N7K-1(config-if-range)#channel-group 1 mode active
```

```
N7K-1(config)#interface port-channel 1
```

```
N7K-1(config-if)#switchport mode trunk
```

```
N7K-1(config-if)#switchport trunk allowed vlan 1-200
```

```
N7K-1(config-if)#spanning-tree port type network
```

N7K-2和N7K-1配置完全相同, 此处以及往后均略

2: 配置两台N7K与N5K的互联线为普通的以太通道

N7K:

```
N7K-1(config)# interface ethernet 1/23
```

```
N7K-1(config-if-range)#switchport
```

```
N7K-1(config-if-range)#channel-group 2 mode active
```

```
N7K-1(config)#interface port-channel 2
```

```
N7K-1(config-if)#switchport mode trunk
```

```
N7K-1(config-if)#switchport trunk allowed vlan 1-200
```

```
N7K-1(config-if)#spanning-tree port type network
```

N5K:

```
N5K(config)# interface ethernet 1/19,ethernet 1/20
```

```
N5K(config-if-range)#switchport
```

```
N5K(config-if-range)#channel-group 2 mode active
```

```
N5K(config)#interface port-channel 2
```

```
N5K(config-if)#switchport mode trunk
```

```
N5K(config-if)#switchport trunk allowed vlan 1-200
```

```
N5K(config-if)#spanning-tree port type network
```

3: 两个交换器设定统一的vpc domain ID

```
N7K-1 (config)#feature vpc
```

```
N7K-1 (config)#vpc domain 10
```

4: 设定Peer keepalive link,并确认启用。

```
N7K-1 ( config-vpc-domain)#peer-keepalive destination 192.168.1.2 source  
192.1.68.1.1 vrf management 或者其他VRF 比如: VPC-KEEPALIVE
```

5: 设定peer-link,并确认启用。

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

```
N7K-1(config)#interface port-channel 1
N7K-1(config-if)#peer-link
6: 可以启用Peer-gateway 和Peer-switch 来优化 VPC
N7K-1 (config)#vpc domain 10
N7K-1 (config-vpc-domain)#Peer-gateway
N7K-1 (config-vpc-domain)#Peer-switch
7: 针对不同的以太通道配置不同的VPC组号。
N7K-1(config)#interface port-channel 1
N7K-1(config-if)#vpc 1
N7K-1(config)#interface port-channel 2
N7K-1(config-if)#vpc 2
```

第二十六节 FEXes---FEXes 在 NX 系列中的支持

N5000+N2000组成了一台集中式结构的虚拟交换机，集中式是指所有的流量都要经过N5000交互，N2000不提供本地交换能力，只是作为N5000的接口扩展。对应802.1Qbh结构，N5000就是CB，而N2000就是PE。组合出来的虚拟交换机中，N5000就是带转发芯片和交换芯片的主控板，而N2000则是接口板，整体更像Cisco早期的4500系列机框或使用主控板PFC进行转发的6500系列机框，但是在N5000盒子内部又是以分布式结构处理转发芯片与交换芯片连接布局。

N5000+N2000实现了Cisco的FEX典型结构（Fabric Extend，等同于Port Extend）。在N5000上看到每台N2000就是以一个FEX节点形式出现的接口板。N2000拥有两种物理接口类型，连接下游设备（可以是服务器或N2000，FEX结构支持级联扩展）的HIF（Host Interface）和连接上游N5000和N2000的NIF（Network Interface），此两种接口是固定于面板上的，且角色不可变更。以2248T举例，右侧黄色接口为NIF，其他为HIF。



FEXes互联模式总的来说有三种：

- 1:straight-through using static pinning
- 2:straight-through using dynamic pinning
- 3:Active-active FEX using vpc

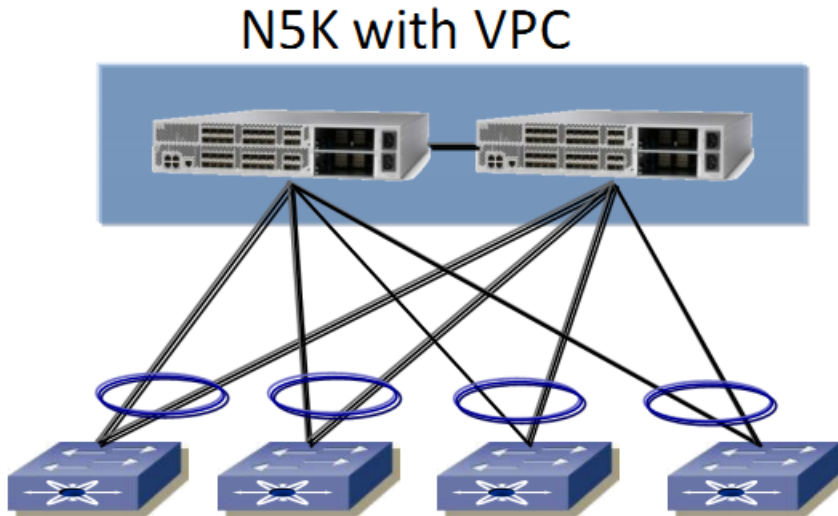
N7K目前只支持straight-through using static pinning

N5K支持全部三种

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



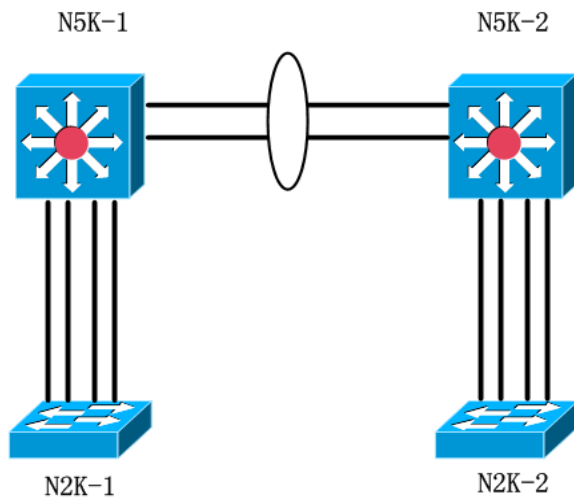
Standard Switch with Port-Channel Support

第二十七节 static pinning

straight-through using static pinning:此种模式要求N2000中，每一组端口都具有自己专用的上行转发线路，并且多条上行线路之间相互不能互相转换。

举例：一块48口N2000，需要4条10GE口与一台N5K相连，每条10GE口顺序对应连续的12个1GE端口，当然其中任何一条中断时，其所对应的12个端口将显示故障。

注意N2K有专门的4个上联10GE口，而N5K则无此限制。



```
N5K-1(config)#feature fex
```

```
N5K-1(config)# fex 111
```

```
N5K-1(config-fex)#description "FEX 111, rack 1,top"
```

```
N5K-1(config-fex)#pinning max-links 4 (4为4条, 如果实际只有2条上联, 那么设定为2)
```

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

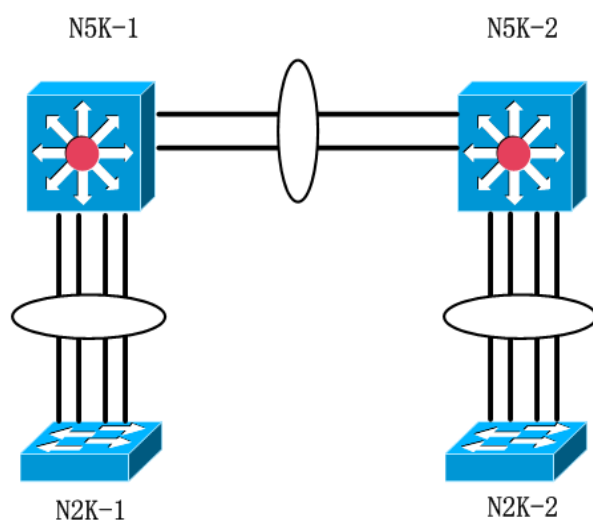
```
N5K-1(config)#interface ethernet 1/1-4
N5K-1(config-if-range)#switchport mode fex-fabric
N5K-1(config-if-range)#fex associate 111
```

```
show fex
show fex detail
```

第二十八节 dynamic pinning

straight-through using dynamic pinning

动态固定模式，通过port-channel，实现如果N2K与N5K的上联线路其中一条失效，并不会影响到N2K的物理端口的状态，只是故障的那条负荷会转移到其余正常的线路上。



```
N5K-1(config)#feature fex
N5K-1(config)# fex 121
N5K-1(config-fex)#description "FEX 121, rack 2,top"
N5K-1(config-fex)#pinning max-links 1 （这个模式只需要1即可）
```

```
N5K-1(config)#interface ethernet 1/9-12
N5K-1(config-if-range)#switchport mode fex-fabric
N5K-1(config-if-range)#channel-group 21
```

```
N5K-1(config)#interface port-channel 21
N5K-1(config-if-range)#fex associate 121
```

本作品仅赠于本文档所有付出与淘宝的朋友







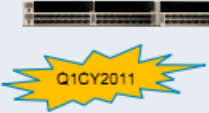

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

第二十九节 FEXes 在 NX 系列中的支持

Nexus 5000 Parent Switch

All Nexus 5000 are VNTag Capable

Parent Switch	FEX Capable	# FEX	Optics/Transceivers Supported
Nexus 5020 		12	Passive CX-1 SFP+ (1/3/5m) Active CX-1 SFP+ (7/10m) SR SFP+ (MMF) – OM3 300m LR SFP+ (SMF) – 300m (FCoE no drop)* FET SFP+ (MMF) – OM2 20m, OM3 100m LRM SFP+ - Not Supported
Nexus 5010 		12	Passive CX-1 SFP+ (1/3/5m) Active CX-1 SFP+ (7/10m) SR SFP+ (MMF) – OM3 300m LR SFP+ (SMF) – 300m (FCoE no drop)* FET SFP+ (MMF) – OM2 20m, OM3 100m LRM SFP+ - Not Supported
Nexus 5548 		16	Passive CX-1 SFP+ (1/3/5m) Active CX-1 SFP+ (7/10m) SR SFP+ (MMF) – OM3 300m LR SFP+ (SMF) – 300m (FCoE no drop)* FET SFP+ (MMF) – OM2 20m, OM3 100m LRM SFP+ - Not Supported
Nexus 5596 (Q1CY11) 		16	Passive CX-1 SFP+ (1/3/5m) Active CX-1 SFP+ (7/10m) SR SFP+ (MMF) – OM3 300m LR SFP+ (SMF) – 300m (FCoE no drop)* FET SFP+ (MMF) – OM2 20m, OM3 100m LRM SFP+ - Not Supported

* Potential to increase supported FEX to Nexus 5500 fabric link distances with 2248/2232 and Future FEX

本作品仅赠于本档所有付出与淘宝的朋友











<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

Nexus 7000 Parent Switch

FEX enabled line cards – VNTag Capable



N7K Parent Switch I/O Module	FEX / vnTag support	Optics/Transceivers supported on N7K + N2K combination
N7K-M132XP-12 32-port SFP+ M1 I/O module \$70,000 		Passive CX-1 SFP+ (1m/3m/5m): No Active CX-1 SFP+ (7m/10m): Yes SR SFP+ (MMF): Yes – OM1 26m OM3 300m LR SFP+ (SMF): Yes – up to 10km LRM SFP+: No FET SFP+ (MMF): Yes – OM2 25m OM3 100m
N7K-M108X2-12L 8-port X2 M1 I/O module \$44,000 		<i>FEX not supported</i>
N7K-M132XP-12L 32-port SFP+ M1 XL I/O module \$70,000 		Passive CX-1 SFP+ (1m/3m/5m): Yes Active CX-1 SFP+ (7m/10m): Yes SR SFP+ (MMF): Yes – OM1 26m OM3 300m LR SFP+ (SMF): Yes – up to 10km LRM SFP+: No FET SFP+ (MMF): Yes – OM2 25m OM3 100m
N7K-F132XP-15 32-port D1 I/O module \$35,000 		<i>FEX not supported</i>
N7K-F2 series 		Passive CX-1 SFP+ (1m/3m/5m): Yes Active CX-1 SFP+ (7m/10m): Yes SR SFP+ (MMF): Yes – OM1 26m OM3 300m LR SFP+ (SMF): Yes – up to 10km LRM SFP+: No FET SFP+ (MMF): Yes – OM2 25m OM3 100m

Nexus 5000 and Nexus 7000 Parent Switch

FEX Support

	N5K Support	N7K Support
Number of FEX supported	16	32
N2K-C2148T-1GE	✓	No
N2K-C2248TP-1GE	✓	✓
N2K-C2224TP-1GE	✓	Q1CY11
N2K-C2232PP-10GE	✓	Q1CY11
N2K-C2232TP-10GE	Target Q1CY11	Target Q1CY11

本作品仅赠于本档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

Nexus 2148T, 2248TP, 2232PP Capabilities

Features	2148	2248	2232
Fabric Ports (NIF)	4	4	8
Fabric Link Port Speed (NIF)	10Gbps	10Gbps	10Gbps
Port Channels on Fabric Links (NIF)	1 x 4 ports maximum Hash L2/L3 fields	1 x 4 ports maximum Hash L2/L3/L4	1 x 8 ports maximum Hash L2/L3/L4
Host Ports (HIF)	48	48	32
Host Port Speeds (HIF)	1 Gbps only	100Mb/1Gbps	1Gbps/10Gbps (No 1 Gbps at FCS, 2HCY10)
Local Port Channels on Host Ports (HIF)	Not Supported	Max 8 ports per port channel, Max of 24 port channels per 2248 Hash L2/L3/L4	Max 8 ports per port channel, Max of 16 port channels per 2232 Hash L2/L3/L4
FCoE/DCB Servers	No	No	Yes
Supported Parent Switches	Nexus 5010/5020/5548 Nexus 5596 - Q1CY11	Nexus 5010/5020/5548 Nexus 7000 M1 Nexus 5596 - Q1CY11 Nexus 7000 F2 - 2HCY11	Nexus 5010/5020/5548 Nexus 5596 - Q1CY11 Nexus 7000 M1 - Q1CY11 Nexus 7000 F2 - 2HCY11

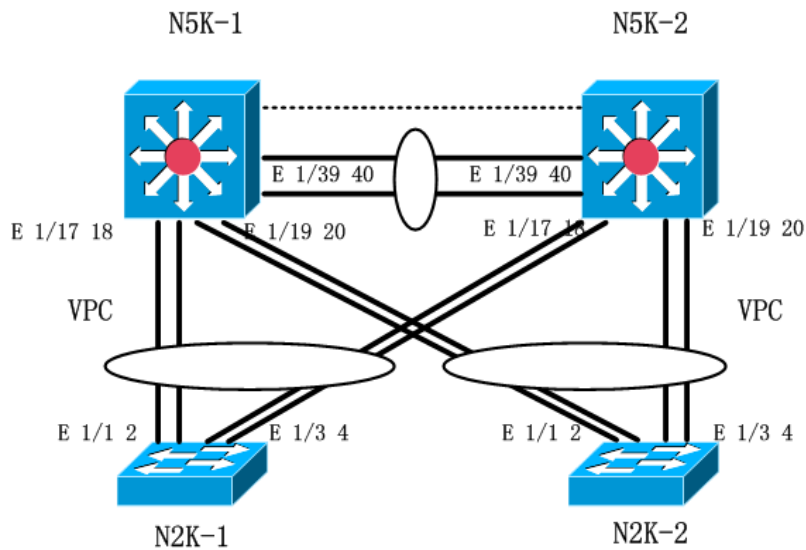
第三十节 active-active FEX

Active-active FEX using vpc

这种模式相对复杂，N2K在这种模式中可以被两台N5K同时管理，但是N5K们对应的配置必须一致，同时VPC必须被设置来组合FEX上行连接成单一的portchannel.

因为这种模式在N2K和N5K之间是基于VPC的portchannel，所以动态固定（dynamic pinning)自动被使用。

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



在N5K-1, N5K-2上启用FEX特性

```
N5K-1(config)#feature fex
```

```
N5K-1(config)# fex 131
```

```
N5K-1(config-fex)#description "FEX 131, rack 3,top"
```

```
N5K-1(config-fex)#pinning max-links 1 (这个模式只需要1即可)
```

将N5K上用于下联N2K的4个接口指定FEX接口模式以及划入以太通道31

```
N5K-1(config)#interface ethernet 1/17-20
```

```
N5K-1(config-if-range)#switchport mode fex-fabric
```

```
N5K-1(config-if-range)#channel-group 31
```

启用VPC特性, 统一VPC域, 并且指定对方的管理地址作为peer-keepalive link

```
N5K-1 (config)#feature vpc
```

```
N5K-1 (config)#vpc domain 37
```

```
N5K-1 (config-vpc-domain)#peer-keepalive destination 192.168.1.2
```

将N5K-1和N5K-2之间互联的2条线路划入以太通道1

```
N5K-1(config)#interface ethernet 1/39-40
```

```
N5K-1(config-if-range)#channel-group 1
```

将以太通道1配置为trunk模式, 并启用peer-link来让N5K之间同步信息。

```
N5K-1(config)#interface port-channel 1
```

```
N5K-1(config-if)#switchport mode trunk
```

```
N5K-1(config-if)#vpc peer-link
```

将N5K-1和N5K-2的以太通道31划入一致的VPC组, 并配置FEX模式。

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

```
N5K-1(config)#interface port-channel 31
N5K-1(config-if)#vpc 31
N5K-1(config-if)#fex associate 131
```

FEX配置完毕后，当一台N5K对应的接口配置改变时，对方N5K也要进行更改：

```
N5K-1(config)#interface ethernet 131/1/1
N5K-1(config-if)#switchport access vlan 10
```

```
N5K-2(config)#interface ethernet 131/1/1
N5K-2(config-if)#switchport access vlan 10
```

```
show vpc consistency-parameters interface ethernet 131/1/2
```

第三十一节 N7K 中的 FEX

N7K中的FEX和N5K不太相同，因N7K为VDC的原因，需要使用专用指令开启FEX特性，但是你可以在指定的VDC中关闭FEX特性。N7K中强制使用port-channel关联N2K

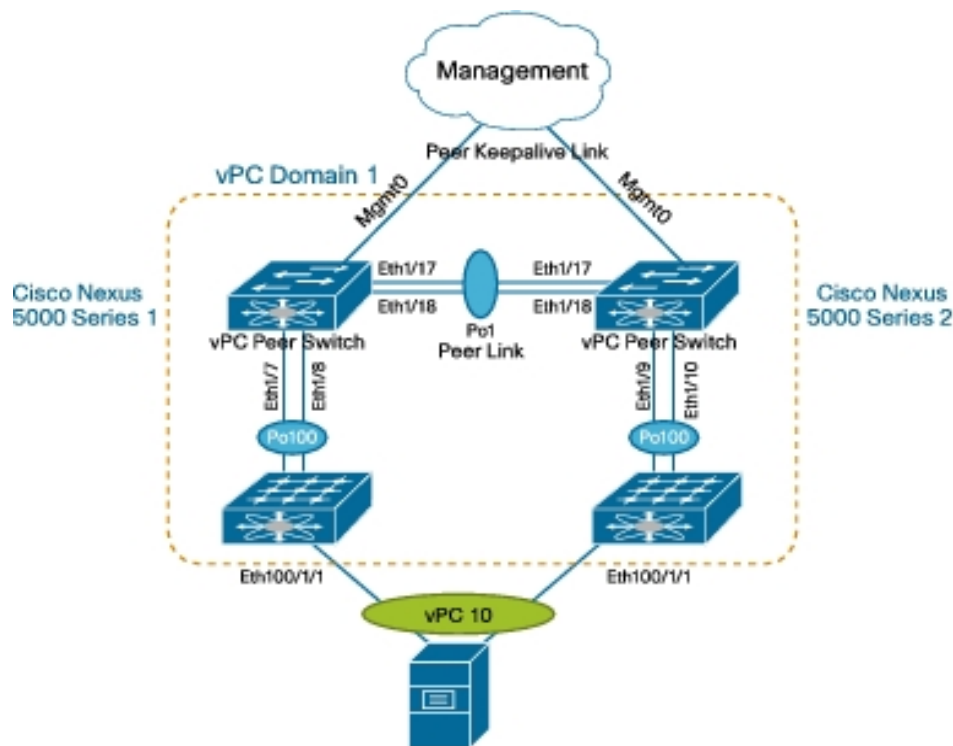
```
N7K-1 (config)#install feature-set fex
N7K-1 (config-vdc)#no allow feature-set fex
```

配置举例：

```
N7K-1 (config)#install feature-set fex
N7K-1(config)#feature-set fex
N7K-1(config)# fex 141
N7K-1(config-fex)#description "FEX 141, rack 4,top"
N7K-1(config-fex)#pinning max-links 4 (N7K中，这条不必再打)
N7K-1(config)#interface ethernet 1/1-2, ethernet 1/9-10
N7K-1(config-if-range)#switchport
N7K-1(config-if-range)#switchport mode fex-fabric
N7K-1(config-if-range)#channel-group 41
N7K-1(config-if-range)#no shutdown
N7K-1(config)#interface port-channel 41
N7K-1(config-if-range)#fex associate 141
```

第三十二节 FEX 在服务器中的应用举例

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



上图的口字型互联只是举例，N2K和N5K也可采用full mesh互联。

下面的配置仅为采用口字型互联的例子，full mesh结构中，其他都一样，只需要稍微调整一下N2K的4个上联口即可。

注释：该例子中，N5K和N2K先做一个dynamic pinning 的VPC，然后将N2K中连接服务器的端口划入VPC组，加入port-channel,然后在服务器的双网卡上也配置为普通的port-channel即可。

Step 1. Configure the management interface IP address and default route.

```
N5k-1(config)# int mgmt 0
N5k-1(config-if)# ip address 172.25.182.51/24
N5k-1(config-if)# vrf context management
N5k-1(config-vrf)# ip route 0.0.0.0/0 172.25.182.1
Step 2. Enable vPC and LACP.
```

```
N5k-1(config)# feature vpc
N5k-1(config)# feature lacp
Step 3. Create a VLAN.
```

```
N5k-1(config)#vlan 101
Step 4. Create the vPC domain.
```

```
N5k-1(config)# vpc domain 1
Step 5. Configure the vPC role priority (optional).
```

本作品仅赠于本档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

```
N5k-1(config-vpc-domain)# role priority 1000
```

Step 6. Configure the peer keepalive link. The management interface IP address for Cisco Nexus 5000 Series Switch 2 is 172.25.182.52.

```
N5k-1(config-vpc-domain)# peer-keepalive destination 172.25.182.52
```

Note:

-----:: Management VRF will be used as the default VRF ::-----

Step 7. Configure the vPC peer link. Note that, as for a regular interswitch trunk, trunking must be turned on for the VLANs to which the vPC member port belongs.

```
N5k-1(config-vpc-domain)# int ethernet 1/17-18
```

```
N5k-1(config-if-range)# channel-group 1 mode active
```

```
N5k-1(config-if-range)# int po1
```

```
N5k-1(config-if)# vpc peer-link
```

```
N5k-1(config-if)# switchport mode trunk
```

```
N5k-1(config-if)# switchport trunk allowed vlan 1,101
```

Step 8. Configure the Cisco Nexus 2000 Series Fabric Extenders and the fabric interface.

```
N5k-1(config)# fex 100
```

```
N5k-1(config-fex)# pinning max-links 1
```

Change in Max-links will cause traffic disruption.

```
N5k-1(config-fex)# int e1/7-8
```

```
N5k-1(config-if-range)# channel-group 100
```

```
N5k-1(config-if-range)# int po100
```

```
N5k-1(config-if)# switchport mode fex-fabric
```

```
N5k-1(config-if)# fex associate 100
```

Step 9. Move the fabric extender interface to vPC. After fabric extender 100 (fex 100) comes online, create the PortChannel for interface eth100/1/1 and move the PortChannel to the vPC. Note that the PortChannel number and vPC number can be different, but the vPC number must be the same on both Cisco Nexus 5000 Series Switches.

```
N5k-1(config-if)# int ethernet 100/1/1
```

```
N5k-1(config-if)# channel-group 10
```

```
N5k-1(config-if)# int port-channel 10
```

```
N5k-1(config-if)# vpc 10
```

```
N5k-1(config-if)# switchport access vlan 101
```

The configuration steps for the second switch, Cisco Nexus 5000 Series Switch 2, are:

```
N5k-2(config)# int mgmt 0
```

```
N5k-2(config-if)# ip address 172.25.182.52/24
```

```
N5k-2(config-if)# vrf context management
```

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

```
N5k-2(config-vrf)# ip route 0.0.0.0/0 172.25.182.1
N5k-2(config)# feature vpc
N5k-2(config)# feature lacp
N5k-2(config)# vlan 101
N5k-2(config)# vpc domain 1
N5k-2(config-vpc-domain)# peer-keepalive destination 172.25.182.51
Note:
```

-----:: Management VRF will be used as the default VRF ::-----

```
N5k-2(config-vpc-domain)# int ethernet 1/17-18
N5k-2(config-if-range)# channel-group 1 mode active
N5k-2(config-if-range)# int po1
N5k-2(config-if)# vpc peer-link
N5k-2(config-if)# switchport mode trunk
N5k-2(config-if)# switchport trunk allowed vlan 1,101
N5k-2(config)# fex 100
N5k-2(config-fex)# pinning max-links 1
Change in Max-links will cause traffic disruption.
N5k-2(config-fex)# int e1/9-10
N5k-2(config-if-range)# channel-group 100
N5k-2(config-if-range)# int po100
N5k-2(config-if)# switchport mode fex-fabric
N5k-2(config-if)# fex associate 100
N5k-2(config-if)# int ethernet 100/1/1
N5k-2(config-if)# channel-group 10
N5k-2(config-if)# int port-channel 10
N5k-2(config-if)# vpc 10
```

```
N5k-2(config-if)# switchport access vlan 101
```

For the deployment scenario in Figure 5, the fabric extender is dual-connected to a pair of Cisco Nexus 5000 Series Switches. Most vPC-related configuration steps are the same as in the previous example, except that the fabric interfaces on the Cisco Nexus 5000 Series Switches will be moved to the vPC rather than to the fabric extender host interface.

第三十三节 路由

N7K配置路由，都要L3的许可，除非你只需要静态路由和RIP V2、HSRP或者阉割版的OSPF（256条路由）那么就可以不买企业的LIC，使用默认的。

现在OSPF不再用 10的8次方/BW了，用 4 X 10的10次方/BW，ISIS也是。

OSPF的network的命令宣告模式取消，不支持，只能在接口指定。

eigrp: 在N7K中，router eigrp 90中 90可以是进程号码，也可以是AS号码，如果有指明命令：

autonomous-system 200，那么90就变进程号码，没有特别指明，就是AS号码。

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

在普通IOS中，启用VRF的话，90变进程号码。

MPLS VRF方面和普通的IOS创建差不多

N7K中，PBR也要特别启用一下。

IPV6不用特别启动，地址直接在接口下配置。

OSPF V3对应IPV6 V2对应 IP V4

ISIS只是N7K支持

-

第三十四节 OSPF

OSPF IPV4:

```
N7K-1(config)#feature ospf(需要企业版LIC)
```

```
N7K-1(config)#router ospf 1
```

```
N7K-1(config-router)#router-id 10.10.10.10
```

```
N7K-1(config-router)#log-adjacency-changes
```

```
N7K-1(config-router)#auto-cost reference-bandwidth 100 Gbps (指定参考带宽，NX-OS默认40G)
```

```
N7K-1(config)#interface vlan 10,vlan 20-25
```

```
N7K-1(config-if-range)#ip router ospf 1 area 11
```

```
N7K-1(config-if-range)#ip ospf passive-interface
```

```
N7K-1(config)#interface ethernet 1/12-15
```

```
N7K-1(config-if-range)#ip router ospf 1 area 0
```

```
N7K-1(config-if-range)#ip ospf authentication message-digest
```

```
N7K-1(config-if-range)#ip ospf message-digest-key 1 md5 S3cr3t
```

OSPF IPV6:

```
N7K-1(config)#feature ospfv3(需要企业版LIC)
```

```
N7K-1(config)#router ospfv3 100
```

```
N7K-1(config-router)#router-id 10.10.10.10
```

```
N7K-1(config)#interface vlan 10,vlan 20-25
```

```
N7K-1(config-if-range)#ipv6 router ospfv3 100 area 11
```

第三十五节 EIGRP

EIGRP IPV4:

```
N7K-1(config)#feature eigrp(需要企业版LIC)
```

```
N7K-1(config)#router eigrp 1
```

```
N7K-1(config-router)#router-id 10.10.10.10
```

```
N7K-1(config-router)#log-adjacency-changes
```

```
N7K-1(config-router)#auto-cost reference-bandwidth 100 Gbps (指定参考带宽，NX-OS默认40G)
```

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

```
N7K-1(config)#interface vlan 10,vlan 20-25
N7K-1(config-if-range)#ip router eigrp 1
N7K-1(config-if-range)#ip passive-interface eigrp
N7K-1(config)#interface ethernet 1/12-15
N7K-1(config-if-range)#ip router eigrp 1
N7K-1(config-if-range)#ip authentication mode eigrp 1 md5
N7K-1(config-if-range)#ip authentication key-chain eigrp 1 EIGRP-CHAIN
N7K-1(config)#key chain EIGRP-CHAIN
N7K-1(config-keychain)#key 1
N7K-1(config-keychain)#key-string S3cr3t
```

EIGRP IPV6:

```
N7K-1(config)#feature eigrp(需要企业版LIC)
N7K-1(config)#router eigrp 200
N7K-1(config-router)#router-id 10.10.10.10
N7K-1(config-router)#address-family ipv6 unicast
N7K-1(config)#interface vlan 10,vlan 20-25
N7K-1(config-if-range)#ipv6 router eigrp 200
```

第三十六节 ISIS

```
N7K-1(config)#feature isis(需要企业版LIC)
N7K-1(config)#router isis DC
N7K-1(config-router)#net 49.0001.1921.6801.1011.00
N7K-1(config-router)#is-type level-1 (预设是level-1-2)
N7K-1(config-router)#log-adjacency-changes
N7K-1(config-router)#reference-bandwidth 100 Gbps (指定参考带宽, NX-OS默认40G)
N7K-1(config)#interface vlan 10,vlan 20-25
N7K-1(config-if-range)#ip router isis DC
N7K-1(config-if-range)#isis passive level-1
N7K-1(config)#interface ethernet 1/12-15
N7K-1(config-if-range)#ip router isis DC
N7K-1(config-if-range)#isis authentication-type md5 level-1
N7K-1(config-if-range)#ip authentication key-chain ISIS-CHAIN level-1
N7K-1(config)#key chain ISIS-CHAIN
N7K-1(config-keychain)#key 1
N7K-1(config-keychain)#key-string S3cr3t
```

第三十七节 BGP

BGP IPV4:

```
N7K-1(config)#feature bgp(需要企业版LIC)
```

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

```
N7K-1(config)#router bgp 65000
N7K-1(config-router)#router-id 10.10.10.10
N7K-1(config-router)#address-family ipv4 unicast
N7K-1(config-router-af)#network 192.168.16.0/20
```

```
N7K-1(config-router)#neighbor 10.1.1.2 remote as 65001
N7K-1(config-router-neighbor)#description ISP peer router
N7K-1(config-router-neighbor)#address-family ipv4 unicast
N7K-1(config-router-neighbor-af)#next-hop-self
```

```
N7K-1(config-router)#neighbor 192.168.16.2 remote as 65000
N7K-1(config-router-neighbor)#description Internal peer N7K-2
N7K-1(config-router-neighbor)#update-source loopback 0
N7K-1(config-router)#address-family ipv4 unicast
```

BGP IPV6:

```
N7K-1(config)#feature bgp(需要企业版LIC)
N7K-1(config)#router bgp 65000
N7K-1(config-router)#router-id 10.10.10.10
N7K-1(config-router)#address-family ipv6 unicast
N7K-1(config-router-af)#network 2001:db8::/32
```

```
N7K-1(config-router)#neighbor 2001:db8:1::1/32 remote as 65001
N7K-1(config-router-neighbor)#address-family ipv6 unicast
```

第三十八节 IPV6 基础配置

```
N7K-1(config)#interface vlan 10
N7K-1(config-if)#ipv6 address 2001:db8:1:10::/64 eui64(前缀64个位必须被指定)
N7K-1(config)#interface ethernet 1/1
N7K-1(config-if)#ipv6 address 2001:db8:ffff:ffff::5/126
N7K-1(config)#interface ethernet 1/2
N7K-1(config-if)#ipv6 address use-link-local-only
N7K-1(config)#interface mgmt0
N7K-1(config-if)#ipv6 address 2001:db8:100:100::100/64
N7K-1(config)#ipv6 route ::/0 2001:db8:ffff:ffff::6
eui64: 前缀64个位必须被指定
use-link-local-only: 想启用IPV6, 但不使用全域性路由IPV6地址。
```

第三十九节 组播

NX-OS系列中不再有DM模式 , 只有稀疏模式
NX-OS只支持 IGMP V2 V3版本

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

NX-OS 组播所支持的三种模式:

1: PIM sparse mode: 这种模式常用于ASM部署。使用一个中心点 (RP), 所有接受者都共有一棵由叶路由器 (直连组播源的路由器) 到RP的共享树。然后由RP再向接受者进行细节路径转发。

2: PIM Bidir: PIM双向模式通常被部署与多对多的多点传送应用。PIM Bidir 只使用一棵植根于RP的共享树。可以让数据流从来源往树的上方流到RP, 以及从RP往树的下方流到接受者。

3: PIM SSM: PIM SSM通常使用于SSM多点传送部署。虽然PIM Bidir只使用一棵共享树, 而PIM SSM则完全相反, 它只使用直接从附挂于组播源的叶路由器所建立的来源树, 返回到来源, 这需要接受者告知来源它有意使用IGMP V3。

(简单的说, PIM SSM模式里不再用RP, 组播源服务器对应的第一个路由器作为树根)

CISCO建议NX-OS环境中, 使用IGMP V3以及 PIM SSM。

不同组播域互联之前主要用MSDP, NX-OS中主要使用MBGP

第四十节 --PIM

```
N7K-1(config)#feature pim
N7K-1(config)#ip pim log-neighbor-changes
N7K-1(config)#ip pim rp-address 192.168.1.1
N7K-1(config)#interface vlan 10
N7K-1(config-if)#ip pim sparse-mode
如何在VRF中指定RP:
N7K-1(config)#vrf context VIDEO
N7K-1(config-vrf)#ip pim rp-address 10.1.1.1
```

第四十一节 IGMP

IGMP不用特别启用, 当某个PIM被启用后, IGMP会自动启用。

NX-OS目前自动启用的IGMP默认版本为V2

```
N7K-1(config)#interface vlan 10
N7K-1(config-if)#ip igmp version 3
```

IGMP在NX-OS系列中在VLAN中缺省启用, 也可单独停用。

```
命令: vlan config 10(早期为vlan 10)
      no ip igmp snooping
```

如果整个环境并没有启用IP多点传送路由, NX中可以设定一个L2交换机来执行IGMP询问功能:

```
N7K-1(config)#interface vlan 10
N7K-1(config-if)#ip igmp snooping querier 192.168.37.1
```

第四十二节 MSDP

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

```
N7K-1(config)#feature msdp
N7K-1(config)#interface loopback 1
N7K-1(config-if)#ip address 192.168.1.1/32
N7K-1(config)#ip msdp peer 192.168.1.2 connect-source loopback1
```

第四十三节 管理

N7K系列预设4种登录角色，其中第一个创建的账号默认是管理员，然后后面有只读的角色。不同的VDC默认共享一个物理管理接口，网段是相同的，可以理解为一张网卡，多个同网段的IP地址，给多个虚拟机用。

你也可以使用每一个VDC分配自己的以太网接口来进行管理，那么就每个VDC的管理就完全分离了。

配置举例：

1: 定义一个用户角色，名字为 op

```
N7K-1(config)#role name op
N7K-1(config-role)#
```

2:给这个用户角色授权可使用的命令：

```
N7K-1(config-role)#rule 1 permit read
N7K-1(config-role)#rule 2 permit read-writ feature diagnostics
N7K-1(config-role)#rule 3 permit read-writ feature ping
N7K-1(config-role)#rule 4 permit read-writ feature vlan
```

3: 给这个用户设定密码： username op password cisco

```
show role name op
sho role pending-diff
```

NX-OS可以配置为不允许太普通的登录密码。password check(命令不准确,意思差不多)如果有数台NX-OS系列，最好使用CFS协议来同步所有的 登录帐号相关的配置。(通过一个组播地址来实现同步)

这里用于同步帐号的CFS协议和peerlink没有关联。

SSH 使用V2。 RSA或者DSA选其一。

配置N7K可以通过一个图形工具进行管理：DCNM，但是配置N7K最好还是使用最全的CLI DCNM只有最基本的LIC,高级的需要买，当你需要用DCNM来管理N7K的时候。

smart call home 你可以设定email地址，发生故障会将一些信息发到邮箱。

支持每天自动将配置保存到一个FTP服务器，需要特别配置

N7K提供特别的XML，可以利用XML中的API进行开发。

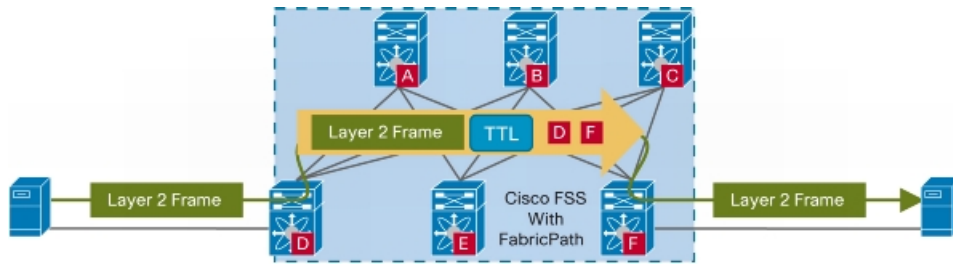
第四十四节 fabric path

fabric path 等于是L2 over L3 ， 需要买对应模块和LIC (可以试用120天)，fabric path是针对目前的L2环境进行的一种优化。(不是连接光存储)

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

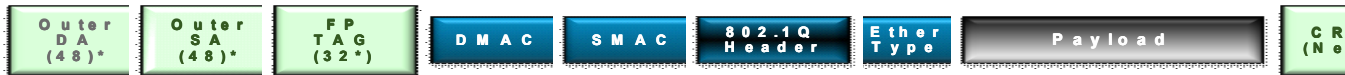
victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



普通以太网帧：



FP帧：使用16字节的头部来建立层次化的地址空间，从而实现增强的特性



不再使用STP，使用ISIS的架构。现在最多16条并行路径，以后可以支持到256条。

也就是说，让L2实现类似L3的多路径负载，只要到达目标距离的COST一样，则负载。

N7K系列交换机fabric path中，学习MAC和普通的不同，N7K需要有人回复ARP的响应，才会将MAC放进MAC表。这样可以防止MAC表溢出攻击。

fabric path的ISIS是自动配置的

要启用fabric path，中间所有的设备都要配置为fabric path功能。一定要F模块。M模块不支持fabric path

如果要使用M模块，除非你的VLAN不需要FP来承载，但是你本地交换机之间是可以相互通信，但是属于M模块的VLAN不能通过FP去往远方，而F模块的可以。

F模块中，要特别配置那些VLAN要走FP，走FP的称为FP VLAN，默认是不走FP的CE VLAN。

配置举例：

```
N7K(config)# feature-set fabricpath
N7K(config)# vlan 10-19
N7K(config-vlan)# mode fabricpath
N7K(config)# interface port-channel 1
N7K(config-if)# switchport mode fabricpath
```

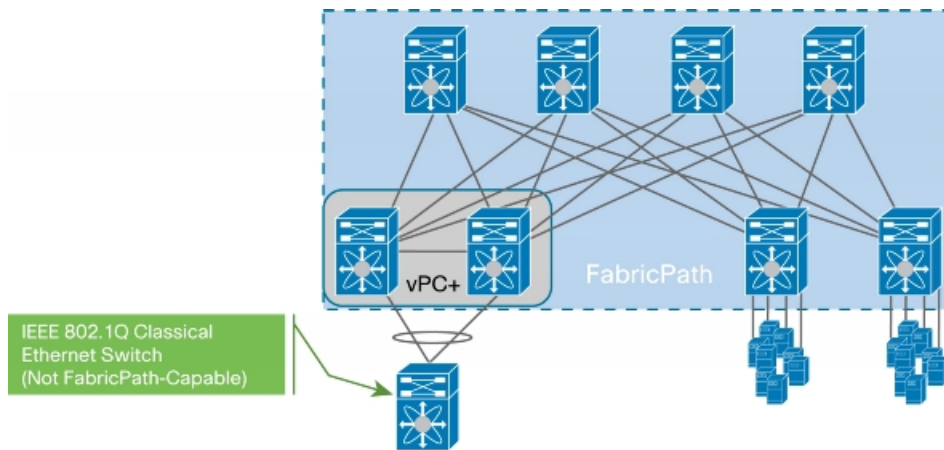
fabric path目前只能用于N7000和N5500（思科目前只推荐使用N7K），而N5000不可以（加模块应该可以）。

VPC中也需要配置一句fabric path的命令，来实现VPC+

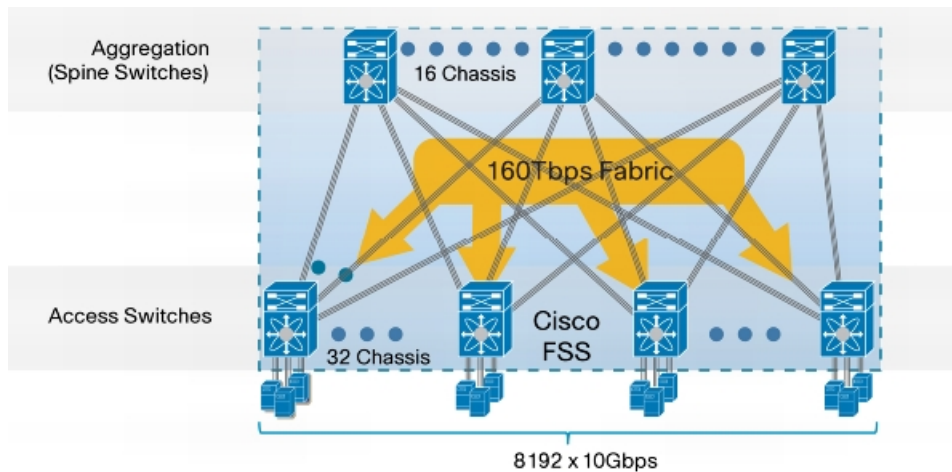
本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



建议N7K配置为根桥或者优先级的数值降低。



第四十五节 DCNM

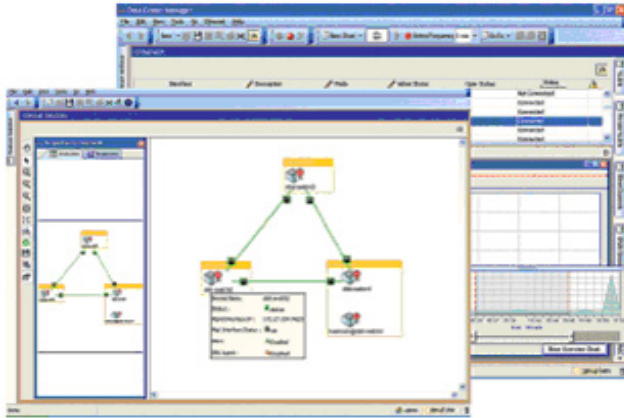
Cisco DCNM是一种全面的、集中的管理解决方案，专门用于数据中心网络的运营（如图1所示）。思科DCNM提供下列重要的优势：

- 1: 多协议感知：管理以太网、IP和
- 2: 故障、配置、记账、性能和安全（FCAPS）支持：提供全面的网络服务生命周期管理支持，强化配置、性能和保障，实现简化的生命周期管理。
- 3: 开放式应用：提供一个可向第三方应用提供状态化网络信息的中间件 API；支持互操作性

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



DCNM有些类似CISCO WORKS，但是功能更加强大：

针对普通的网络：

- 1: 可以显示L1和L2TOP图
- 2: 接收设备故障告警
- 3: 监控设备硬件状态、温度、CPU和内存占用率
- 4: 流量统计等等。
- 5: AAA、RBAC（基于角色的控制）

特别针对DC环境：

- 1: 详细显示每个VDC的实际情况，即时告警
- 2: 监控每个N系列设备软件镜像情况
- 3: 监控和配置各个VPC情况
- 4: 设定与查看fabric path运行状况
- 5: 使用NX-OS自带的配置回复功能，归档历史数据

第四十六节 L2, L3 可靠性

L2的可靠性: BPDU UDLD 链路聚合

L3: NSF（在OSPF等协议配置graceful-restart） HSRP VRRP 等

VRRP和HSRP的区别: VRRP可以将虚地址设定为其中一个路由器已有的本身接口地址, HSRP不可以。HSRP 在多个成员情况下, 只有一个主, 一个备, 但是VRRP是一个主, 多个备。

GLBP: 所有成员可以同时转发, 因为有多MAC地址, 轮询提供给多个客户, 实现负载均衡。

第四十七节 Port-Profile

当大量的接口需要分配相同的参数时,Port-Profile是非常有用的。

```
n7000(config)# port-profile type ?
```

```
  ethernet      Ethernet type
```

```
  interface-vlan Interface-vlan type
```

```
  port-channel  Port-channel type
```

```
n7000(config)# port-profile type ethernet Email-Template
```

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

```
n7000(config-ppm)# switchport
n7000(config-ppm)# switchport access vlan 10
n7000(config-ppm)# spanning-tree port type edge
n7000(config-ppm)# no shutdown
n7000(config-ppm)# description Email Server Port
n7000(config-ppm)# state enabled
n7000(config)# interface ethernet 2/1-2
n7000(config-if-range)# inherit port-profile Email-Template
```

第四十八节 ACS

ACS当然也可以用于NX-OS系列。

如果要实现不同的VDC完全验证方面的不同，需要不同的ACS，或者尽量使用不同的组。现在CISCO建议首选radius协议，LDAP也可以。

NX-OS中，AAA可以调用给一个组，将多个ACS server放入该组中，并且N7K可以通过配置自己发现组成员中故障的成员，并自动屏蔽，也可以设置查询间隔，将已经恢复的服务器，放入组。

配置一个ACS集群组：（下面含义为所有服务器必须使用radius协议，服务器顺序将从上到下询问，最多100个集群组）

```
N7K-1(config)#aaa group server radius RADSERVER
N7K-1(config-radius)#server 10.1.1.1 //将多个ACS服务器加入组
N7K-1(config-radius)#server 10.1.1.2
N7K-1(config-radius)#deadtime 30 //设置健康检查时间，故障的
ACS将被T出
N7K-1(config-radius)#use-vrf management
```

进行普通的ACS配置（这个和老ACS配置类似）（默认使用本地数据库）

```
N7K-1(config)#aaa authentication login default group radius
//设定默认的登录方式，这里选的只是radius协议（可以有：服务器集群、radius\tacacs+\LDAP等服务器的命名子集、NX-OS的本地数据库、或者仅限使用者的名字）
```

```
N7K-1(config)#aaa authentication logins console group RADSERVER
//设定CON口登录的方式，这里选择是服务器集群
```

针对ACS服务器的状态进行监控配置（以分钟为单位）：

```
N7K-1(config)#radius-server host 10.1.1.1 test idle-time 20 //设定健康检查时间
间隔)
N7K-1(config)#radius-server host 10.1.1.1 test username testuser //设定检查帐户和
密码
N7K-1(config)#radius-server deadtime 30 //设定服务器判定故
障时间
```

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

进行手动测试的命令：

```
N7K-1#test aaa server radius 10.1.1.1 test test
```

第四十九节 CFS 部分配置同步协议

CFS 主要用于每个 NX-OS 设备配置方面的同步，包括 radius\tacacs+\callhome\ntp\user&administrator roles

distribute命令使得对应的应用组播同步出去。一个人改的时候，其他人无法更改。

CFS使用缺省的组播地址（需要手动启动该CFS特性），IPV4和IPV6,组播地址可以手动更改，同组N7K的

组播地址需要一致。所以需要这个功能的N7K需要启动PIM。（N5K也可以一起同步）

可以使用cfs region x 来将N7K们划分为不同的同步组。

在使用CFS配置的时候，关于CFS中的配置，没做commit之前，你刚才打入之前的配置都不生效

CFS 配置举例：

1: 指定用于配置同步CFS协议所使用的组播地址：

```
N7K-1(config)#cfs ipv4 mcast-address 239.255.XX.XX(XX可以为任意数字)
```

```
are you sure?(y/n) [n] y
```

2: 使能CFS服务

```
N7K-1(config)#cfs ipv4 distribute
```

3: 指定CFS的区域号（意思就是可以有多个CFS）

```
N7K-1(config)#cfs region XX
```

```
N7K-1(config-cfs-region)#role //将role功能指定进CFS进行同步
```

```
are you sure?(y/n) [n] y
```

```
N7K-1(config-cfs-region)#radius //将radius功能指定进CFS进行同步
```

```
are you sure?(y/n) [n] y
```

4: CFS中，使能role和radius同步功能

```
N7K-1(config)#role distribute
```

```
N7K-1(config)#radius distribute
```

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

第五十节 ISSU

ISSU:不中断服务而升级IOS。需要2张主控板，先升级后备主控，再切换到备主控运行网络，再升级

主控。只有一条命令:

```
install all kickstart bootflash:n7000-s1-kickstart.4.0.1a.bin system
bootflash:n7000-s1-dk9.4.0.1a.bin
```

不过这命令之前，要将新的镜像放置在FLASH中。

sys switchover 切换主控

有的时候升级完NX-OS后，有些模块中EPLD也许也要对应升级。

```
install module 5 epld bootflash:n7000-s1-epld.5.1.2.img
```

第五十一节 ACL

NX-OS中只有类似IOS中的扩展ACL，掩码可以使用/24之类的数字代替。

ATOMIC ACL是在改动ACL后，不立即替换原ACL空间，另寻空间存放。

该特性为默认开启，可以关闭这个特性。

NX-OS ACL支持 config session模式，预制配置和应用，但不生效，检查预应用生效后，才可以应用。

NX-OS中也有 PACL (思科低端交换机中有此功能): port ACL放在L2的接口下的

ACL可以延迟生效，以及定义时间范围。(是否支持配置并应用ACL之后，没有最终确认的话，延迟一段时间后，恢复之前ACL配置??)

N7K: SGT (安全群组标签), SGACL意思是根据服务器的分组的标签，来应用不用组的ACL，制定组和组之间的访问控制策略。SGACL通过ACS创建，不同的用户组给不同的组标签。

(ACS版本的需求?? 应该是4.2以后的)

如果是VM在服务器中的话，建议使用MAC地址的验证，或者用VN-LINK来做。

第五十二节 QOS

N7K全部模块提供inb QOS

N7K的QOS都是分开做的，和之前IOS的QOS什么都放在policymap中做不一样。

N7K的policy map只做 策略

N7K分queuing policy map和 QOS policy map

inbound先执行queuing policy map再执行QOS policy map

outbound反之。

queuing class map 已经被指定，无法更改。

1G 2个队列， 10G的8个队列

N5548必须加4条命令 (WHY 思科不默认加。): service-policy 的几条

做了之后，保证给FCOE 4G带宽

N5K和N7K不用，默认做好。

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

第五十三节 CMP

SUP1上的CMP口，效果基本等于带八爪鱼线的2511

CMP有自己的精简IOS，独立于N7K之外，利用CMP访问N7K的主控板。

CMP本身带有自己独立的以太网口。

attach cmp 类似于CATALIST里面的session

sup是放5,6号插槽。配置CMP时要指定哪一个SUP的CMP。

N7K开启服务命令：ssh server enable（默认启动） telnet server enable

CMP要升级，那么2个SUP的CMP模块也都要升级，升级IOS时，安装的时候要特别指明哪个模块。

moniter CP (N7K console 的信息显示在CMP)

attach CP（从CMP进入N7K console)

第五十四节 BFD

BFD：在启用BFD启用之前，必须启用命令：no hardware ip verify address identical
然后启用BFD后，再使用no ip redirects

因为BFD原理是自己发自己IP的包，N7K会认为是一种攻击(N7K自带IDS特性，会查证IP封包头各字段是否异常，预设的实施检查工作之一就是验证封包之来源于目的IP地址是否完全相同)。

```
N7K-1(config)#feature bfd
```

```
N7K-1(config)#router ospf 1
```

```
N7K-1(config-router)#bfd
```

```
N7K-1(config)#router eigrp 1
```

```
N7K-1(config-router)#bfd
```

```
N7K-1(config)#int vlan 10
```

```
N7K-1(config-if)#no ip redirects
```

```
N7K-1(config-if)#hsrp bfd
```

注意：一旦启用BFD，所有互联端口必须no ip redirects

举例：OSPF中启用了BFD特性，那么就要在所有运行OSPF协议的端口上启用no ip redirects

第五十五节 配置回滚

创建检查点配置（默认名称=“auto-#”）您可以每个VDC创建多达10个配置

```
n7000# checkpoint
```

```
Processing the Request... Please Wait
```

```
..... Done
```

检查配置是否被创建为存档

```
n7000# show checkpoint summary
```

```
-----
```

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

Name	UserName	Created at
auto-2 2005	admin	Tue May 24 22:24:01

手动删除的VLAN 20 & 30

```
n7000# config t
```

```
n7000(config)# no vlan 20, 30
```

回滚配置 “auto-2

```
n7000# rollback running-configuration checkpoint auto-2
```

Processing the Request... Please Wait

Generating the Rollbackpatch... Please Wait

Executing the patch... Please Wait

```
`conf t`
```

\\自动将 VLAN 20 & 30

放回running configuration

```
`vlan 20, 30`
```

清除检查点数据库 注意! 此操作将清楚所有检查点配置

```
n7000# clear checkpoint database
```

Processing the Request... Please Wait

..... Done

第五十六节 FCOE

将光存储整合到IP网络，但是仍然采用光存储的光口

第五十七节 SCSI、FC、FCOE 概念

以下概念部分引用自网络：

SCSI: 小型计算机系统接口（英语:Small Computer System Interface; 简写:SCSI），一种用于计算机和智能设备之间（硬盘、软驱、光驱、打印机、扫描仪等）系统级接口的独立处理器标准。SCSI是一种智能的通用接口标准。它是各种计算机与外部设备之间的接口标准。

Fibre Channel: 简称FC，是一种高速网络技术标准（T11），主要应用于SAN（存储局域网）。

光纤和铜制电缆都可以作为FC的传输介质。

从分层协议栈的角度看，FC仅仅包含了从物理层到传输层的规范。它的上层定义了把其他协议作为应用层协议进行封装的接口，如SCSI或IP协议。而将SCSI3封装起来后整个协议，就是FCP（FC Protocol）。

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

FC物理层具有很高的传输带宽，从1Gb/s、2Gb/s、4Gb/s到8Gb/s，采用NMB的编码方式，同步串行方式传输。

FCOE: Fibre Channel over Ethernet 以太网光纤通道。

FCoE是去年4月份由数家IT厂商向美国国家标准协会（ANSI）T11委员会提交的一种新技术标准的提案。

FCoE技术标准可以将光纤通道映射到以太网，从而可以在以太网上传输SAN数据。它能够保护客户在现有FC-SAN上的投资（如FC-SAN的各种工具、员工的培训、已建设的FC-SAN设施及相应的管理架构）的基础上，提供一种以FC存储协议为核心的I/O整合方案。

当前的FCoE技术标准提案可以使用任何速度的网卡，但需要网卡支持802.3x PAUSE机制。

第五十八节 SCSI\FC 传输机制

Fibre channel\SCSI bus和FCOE都是用来传递SCSI协议的。

和普通TCP/IP环境不同的是：fibre channel中的网络设备传输会相互协商，发送和结束数据都相互通知，是比较可靠的。

这是因为：

FC中设备之间有login机制，物理连接后就LOGIN，而不像IP里需要传输TCP数据时才三次握手。

FC数据传输时有类似TCP/IP中的window的概念，接收方通过window告知对方我还有空间，你继续发，或者我的缓冲区已经占满，你延迟再发。这种叫 buffer-to-buffer flow control.

但是FC中，是设备和直接设备之间，而不像TCP中的只是主机对主机。

再结合end-to-end control（终点对终点），比如硬盘和服务器的。比如告诉对方我已经准备完毕。

总结下来，FC的传输机制归类如下：

- 1: point to point oriented 设备到直连设备的LOGIN机制，类似TCP连接的建立
- 2: N_port to N_port connection 逻辑节点的通信（比如服务器和远端服务器）
- 3: Flow controlled: hop-by-hop and end-to-end basis 通过逐跳以及终端对终端的来进行流控制。
- 4: acknowledged 确认类和通信，类似于TCP中的ACK报文的作用。
- 5: Multiple connections allowed per device 每个设备允许多个连接点，类似于TCP/UPD中的多点发送。

下图中：

FC交换机互联端口成为 E port,FC交换机连接终端口为：F port,如果在很久以前FC交换机接了一个HUB成为FL Port.

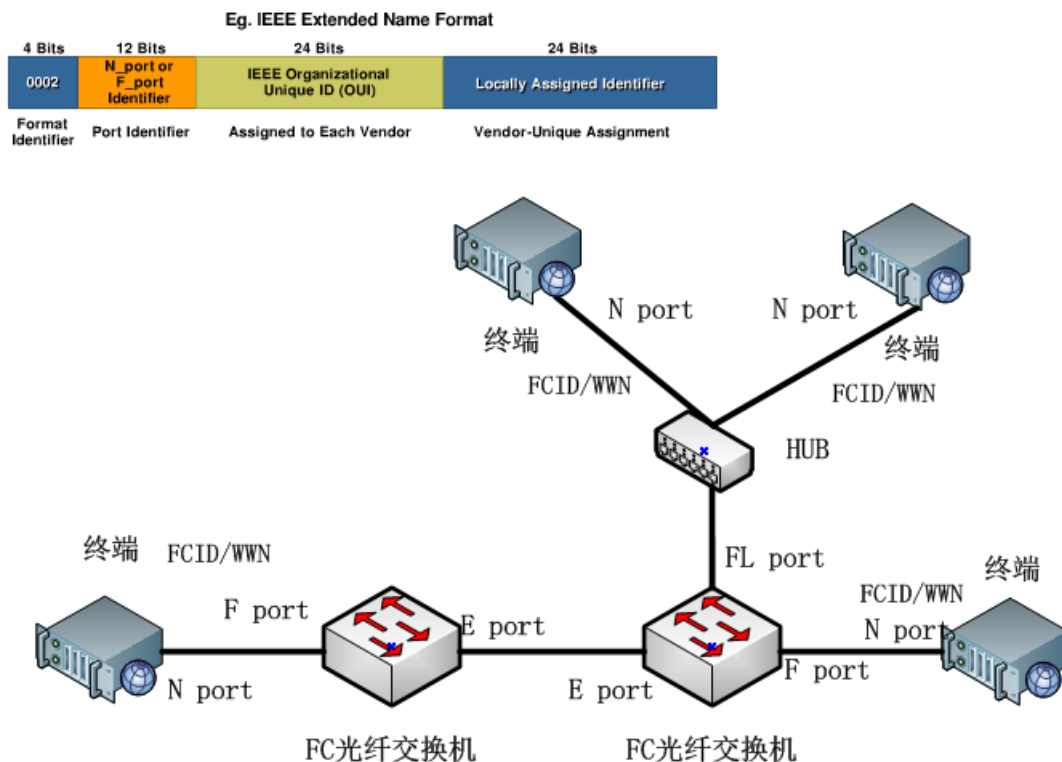
FCID/WWN 类似TCP/IP中的 IP/MAC的关系。

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

每一台FC交换机都将给自己的每一个终端分配独有的FCID。
而FC中的WWN则根据每一个终端的MAC地址创建



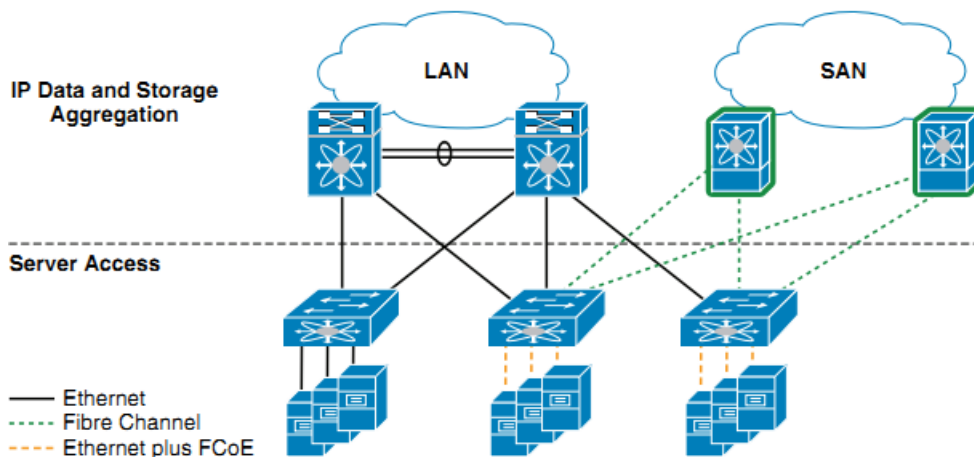
第五十九节 FCOE 机制和配置举例

FCOE顾名思义就是将用于光存储的信号通过以太网来承载，可以简化物理上的布线，统一一套物理架构。

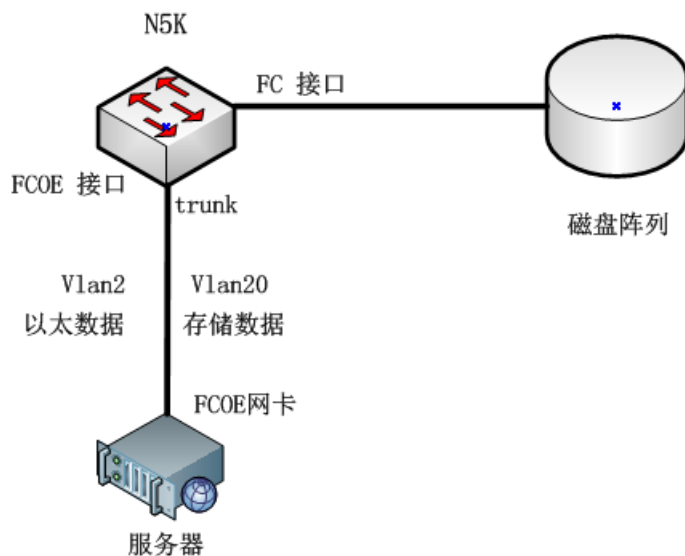
CISCO的FCOE基本原理是通过N5K设备与具备支持FCOE特性网卡的服务器相连接，这块FCOE网卡既跑普通的以太业务数据，也跑光存储数据到磁盘阵列。（注意，同样需要注意的是N5K上用于连接FCOE的端口也要确认为支持FCOE特性）

思科建议的FCOE结合SAN的架构：

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



而实现的方式为将N5K到服务器的互联物理以太线路，封装为trunk协议，逻辑上分为2个VLAN，一个VLAN用于传输以太数据，一个VLAN用于传输存储数据。而N5K和服务器之间在连接时就会有一个LOGIN机制，这个机制让N5K和服务器实现认证、通信。而其中有一个FIP的协议，可以自动分辨哪些走以太通道，哪些走存储通道。以上原理可以根据这个示意图学习。

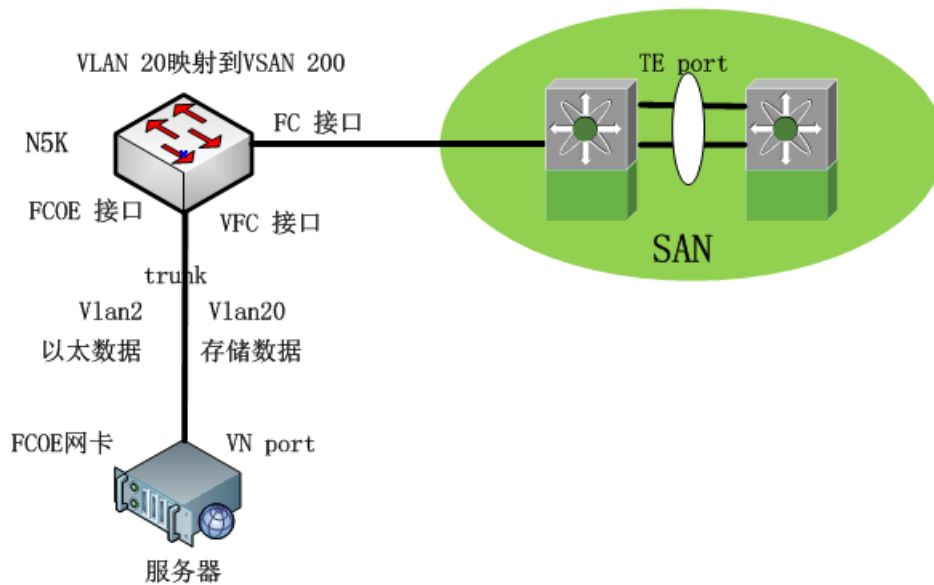


而仅仅是有VLAN的区别是不够的，实际上针对存储的数据有专用的VLAN来传递数据，称为VSAN，因为实际上在NX整个架构中是要模拟整套SAN光存储的架构和协议，所以要在边缘N5K中将普通的以太VLAN和VSAN关联起来。缺省所有的口都放在VSAN1，且4094是默认的不可用VSAN。在思科FCOE架构中，针对原SAN的称谓，有一个大致的规律，就是在SAN的名称前加V。比如：SAN VS VSAN，FC port VS VFC，E port VS VE port 等等。这一段原理可以根据这个示意图学习：

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



个人的建议：思科在设计SAN融合到NX系列中时，设计有些复杂以及有些局限，上图中的SAN环境，如果你创建了多个VSAN，那么在SAN环境中，要需要针对SAN交换机进行特别的配置，要能承载多VSAN，带来一定的麻烦。

实际中，其实大家的光存储架构很单一，基本上服务器和存储都在一个单一个VSAN里面，没有分那么清楚。

所以个人建议设计DC时，最好使用单一VSAN的方式，比如就使用默认的VSAN1，这是VSAN中的本真VLAN，就可以简单的让NX系列和SAN互联即可，就类似普通SW的access接口互联即可，不用去配什么VSAN trunk，TE port什么的，因为光交换方面大家也接触的少，而且也有一定局限性，因为思科在配置的时候都是以自己的MDS光交换机作为环境，实际其他品牌如何弄，未必就这样，大家也未必就清楚。

个人认为，只要遵循简单、稳定、高效、易维护为原则最好。

FCOE配置举例：

1: 在N5K中指定连接服务器的FCOE口，封装为trunk

```
N5K-1(config)# interface ethernet 131/1/1 //131代表N2K的FEX号码，FCOE需要N2232的支持)
```

```
N5K-1(config-if)#switch mode trunk
```

```
N5K-1(config-if)#span port type dege trunk
```

2:指定传递以太数据的VLAN:

```
N5K-1(config-if)#switch trunk native vlan 2(这里只是将本真VLAN修改为2，可以使用默认的vlan1，也可以在服务器里设置使用比如VLAN10来传递数据，不一定一定得用本真VLAN，这里的本真VLAN只用于传递普通以太网数据,也就是说如果不在服务器中设置VLAN，那么以太网数据就走本真VLAN了)
```

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

```
N5K-1(config-if)#switch trunk allowed vlan 2,20
```

3:开启VFC，并将物理接口划入VFC中

```
N5K-1(config)#inter vfc 30  
N5K-1(config-if)#bind inter ethernet 131/1/1  
N5K-1(config-if)#no shut
```

4: 创建普通的VLAN 20和VSAN 的VLAN 200，并将以太VLAN20和VSAN VLAN 200关联起来

```
N5K-1(config)#vlan 20  
N5K-1(config-vlan)#fcoe vsan 200  
N5K-1(config-vlan)#exit  
N5K-1(config)#vsan database  
N5K-1(config-vsan)#vsan 200  
N5K-1(config-vsan)#vsan 200 inter vfc 30 (将VFC关联到VSAN)  
N5K-1(config-vsan)#exit
```

```
show vsan membership  
show vlan fcoe
```

创建VSAN举例:

创建一个VSAN和创建一个VLAN机制类似:

```
N5K-1(config)#vsan databaes  
N5K-1(config-vsan)#vsan 2  
N5K-1(config-vsan)#vsan 2 name AB  
N5K-1(config-vsan)#vsan 3  
N5K-1(config-vsan)#vsan 3 name CD  
N5K-1(config-vsan)#no vsan 2
```

第六十节 FCOE 机制的要点

FCOE的要点:

F PORT接主机，分配主机一个FCID，如果主机由VM虚拟机构成，那么要启用FCIV特性，为多个虚拟机获取多个FCID。

有专门的协议（FIP）来感知，是否要走FCOE（FIP是自动的）。

FCOE卡缺省为一个trunk。

FCOE中，和SAN环境相同，主机需要到SW中进行登录，然后才可以传递存储数据。FCOE专用的MAC就是登录后得到。

现在可以支持多跳FCOE，多个N5K互联的时候，也得是FCOE互联线，并且是V Eport
FCOE也可以做传统的QOS。

本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。

FC中，LLDP，对应CDP，不过是公有标准的邻居发现协议。

要支持FCOE的设备中，就要支持DCB协议，两台设备互联，将会用DCBX来交换信息。DCB本身是一个标准。

一个连接主机的FCOE口，TRUNK，逻辑分为几个虚拟通道：一个是VSAN，一个是本真VLAN 其中VSAN分配一个VLAN，然后将VFC关联到这个VSAN，VSAN VLAN用于传递FCOE数据，

本真VLAN用于传递普通以太网数据。

VFC所在VLAN，要和FC所在的VSAN对应(在VSAN里使用命令来对应)。

注意属于VSAN的VLAN不能和VFC VLAN为同一个，而是在VSAN下关联对应的VFC VLAN。

传数据的VLAN可以自己定义，可以是本真，也可以是其他VLAN，这一点和传统一样。

VFC VLAN不能配为native vlan

缺省所有的口都放在VSAN1，4094是默认的不可用VSAN

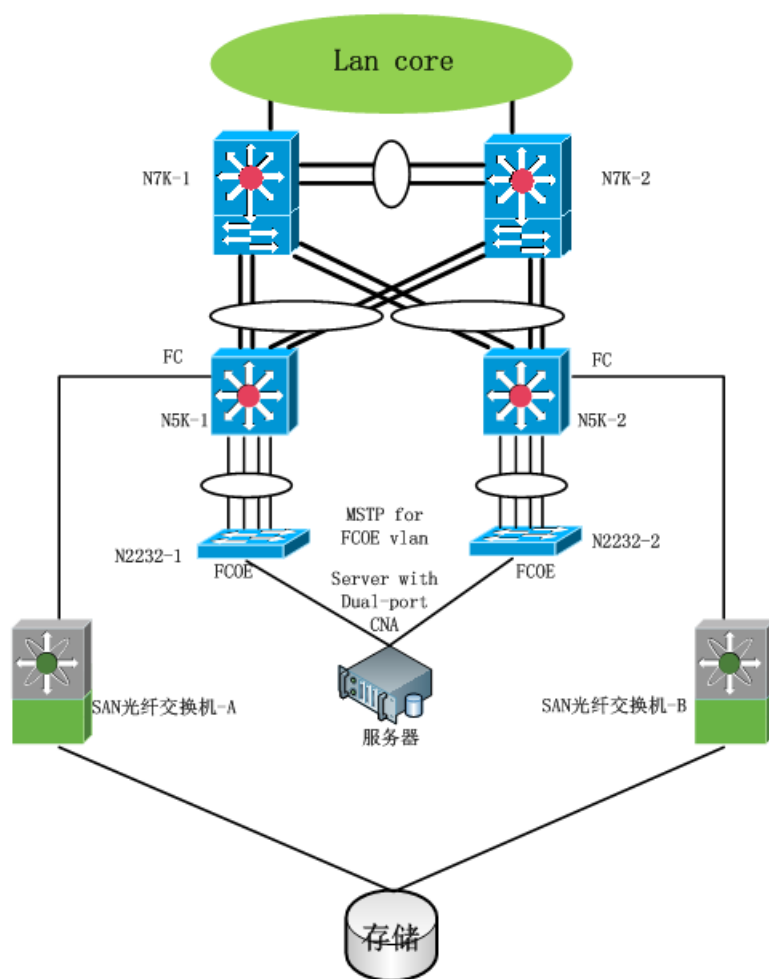
别轻易改动一个交换机的SWID，以免这台SW中所有的FCID都被更改。每一个VSAN都要一个domain ID。

不同的domainID中，可以使用相同的VSAN ID。

第六十一节 个人推荐的 FCOE TOP

以下示意图应该可以满足大部分数据中心的需求：

victory belongs to those that believe in it the most and believe in it the longest 胜利是属于那些最有信心与最有毅力的人。



本作品仅赠于本文档所有付出与淘宝的朋友

<http://ccie.taobao.com/>