彎曲評論

科技 · 人物 · 潮流

思科核心路由器CRS-1的研究（上）

陈怀临，首席科学家
《弯曲评论》
www.tektalk.cn
huailin@tektalk.cn

1. 前言:

在通信领域的高端设备上，核心路由器历来是皇冠上的珍珠。本文对思科的核心路由器CRS-1和CRS-1上的重要部件SPP网络处理器进行考察，从而使得读者能够对CRS-1的体系结构有个整体的了解和把握。

2004年5月25日，思科正式发布其下一代IP核心路由器CRS-1。其原文可参阅：思科CRS-1新闻发布稿。在CRS-1的新闻发布中，关于CRS-1系统的创新和亮点思科是这样描述的：
× Cisco IOS® XR Software, a new member of the Cisco IOS Software family, designed for terabit-scale routing systems built on massively distributed multi-shelf architectures
【笔者译：】思科的IOS-XR操作系统。一个为支持多机箱互联的大规模分布式体系结构而量身定做的一个崭新的IOS家族成员。
× System capacity of up to 92 terabits per second (Tbps)
【笔者译：】系统容量，（在最全配置时），可高达92太比特每秒。（太比特：Terabits。1 Tbps = 1000Gbps)
× Industry's first Optical Carrier (OC)-768c/STM-256c packet interface
【笔者译：】工业界的首次的40G的OC-768光网络接口。（OC-768接口线速为39,813 Mbit/s。通常简称为40G端口。但不要与40G以太端口相混淆。这里是SONET光网络接口。是一种WAN，广域网接入接口。以太端口相当而言是一 个局域网端口的范畴，如现在广泛采用的10GETH以太网端口。OC-XXX指的是数据信号在SONET光纤上传输的速率。）

× Cisco Silicon Packet Processor (SPP), the world's most sophisticated 40-Gbps application-specific integrated circuit (ASIC)
【笔者译：】思科的SPP网络处理器，（在2004年），世界上最复杂，最强大的支持40Gbps的ASIC芯片。
× Extensible Markup Language (XML)-based Cisco Craft Web Interface (CWI), a visual management tool that can manage single-shelf or multishelf systems.
【笔者译：】基于思科的CWIde扩展XML描述语言，一个基于图形界面的管理工具。
× The Cisco Intelligent ServiceFlex design for service flexibility and speed to service
【笔者译：】思科的智能ServiceFlex设计方法，确保服务商能在各个层面提供层2和层3的数据服务。

笔者将会把注意力集中在CRS-1体系结构，CRS-1的SPP网络处理器，和相应的IOS-XR操作系统方面。力图把一个高端核心路由器的整体全貌分析，解释给有兴趣的读者。

此文不允许被转帖，下载和应用在任何商业途径和（或）商业公司研发中；对于教育和非赢利性的目的，可以自由转帖，下载和修改。本文遵守自由软件GNU Free Document License的文档条款。关于GFDL的细节，可参阅GNU站点GFDL条款。

2. 产品系列

如果访问思科的主页，浏览其产品分类，读者可以发现，在路由器这个产品线上，分为Branch，WAN和Service Provider三大类产品。在Service Provider子类里，CRS赫然列为最高端产品。其层次如下图所示。

从图中可见，笔者在前文“思科QuantumFlow处理器与ASR1000的研究”一文中介绍的ASR1000属于边缘路由器（Edge Router）的范畴；而CRS-1是运营服务商（Service

Provider）在主干网上的核心路由器。其地位对于一个公司数据通信产品解决方案是举足轻重，属于重型装备。
CRS-1产品是一个产品家族。其分为2大部分，共4个产品。第一部分叫做线卡机（Line Card Shelf）。第二部分叫做交换连接机（Fabric Card Shelf）。Fabric Card Shelf是用来把多个Line Card Shelf连接成机群的产品。在Line Card Shelf部分，目前思科提供三个产品，分别为4个线卡槽的CRS-1（4 Slot Line Card Shelf），8个线卡槽的CRS-1（8 Slot Line Card Shelf）和16个线卡槽的CRS-1（16 Slot LineCard Shelf)。
下面是CRS-1家族产品系列的图示。

1.CRS-1 4线卡槽单机箱：

思科4线卡槽机箱CRS-1路由器

2. CRS-1 8线卡槽单机箱：

CISCO CRS-1 8-SLOT SINGLE-SHELF

Cisco CRS-1 8-Slot Line Card Chassis

- Eight 40-Gbps line card slots—Using a midplane design, the 8-slot LCC is loaded with 8 MSCs on the back of the chassis, each of which is connected, through the midplane, to an interface module on the front of the chassis.
- 2 dedicated route-processor slots
- 4 dedicated switch fabric card slots (which accommodate 4 switch fabric cards with two switch planes on each)
- Redundant power supplies and fan trays

思科8线卡槽CRS-1路由器

3.CRS-1 16线卡槽单机箱：

CISCO CRS-1 16-SLOT SINGLE-SHELF SYSTEM

Cisco CRS-1 16-Slot Line Card Chassis

- Sixteen 40-Gbps line card slots—Using a midplane design, the 16-slot LCC is loaded with 16 MSCs on the back of the chassis, each of which is connected, through the midplane, to an interface module on the front of the chassis.
- 2 dedicated route-processor slots
- 8 dedicated switch fabric card slots
- 2 dedicated shelf-controller slots
- Redundant power supplies and fan trays

思科16线卡槽CRS-1路由器

4.CRS-1 Fabric交换连接机机箱：

思科Fabric交换连接机箱

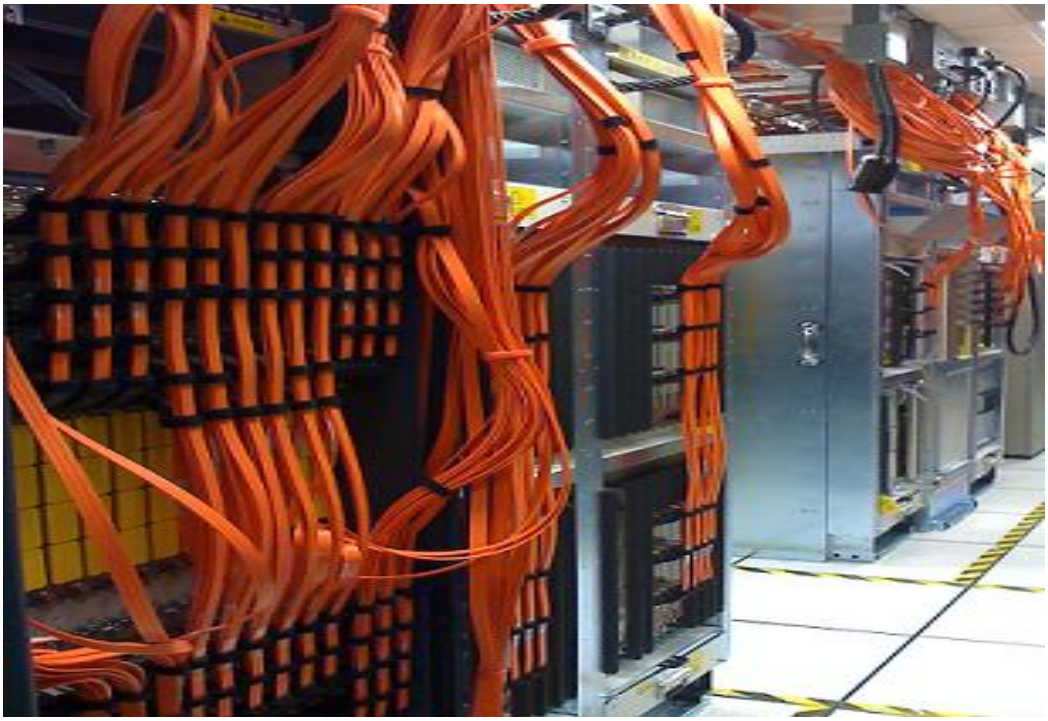
从上图可知，对于4线卡，8线卡，16线卡的CRS-1机器而言，本身就是一个可独立运行的核心路由器。而对Fabric交换机，其本身不是一个路 由器，而是一个非常快速的交换设备从而运营商可以把16线卡的路由器连接成一个路由器群，从而可以形成和支持超大规模的数据通信。读者请注意，4和8线卡 槽的CRS-1不能接入Fabric互联中。

对于4，8和16线卡槽的CRS-1路由器而言，其报文交换吞吐率分别为：（4，8或16）×每个线卡的吞吐率×2.为什么乘2，是因为在Layer-2谈交换吞吐率时，通常是算双工速率（Ingress和Engress）。

因此，4卡CRS-1的交换能力是320Gbps；8卡CRS-1的交换能力是640Gbps；16卡CRS-1的交换能力是1.2Tbps。

对于通过Fabric交换互联形成的机组，CRS-1的交换能力是巨大的。最高配置可以通过8个Fabric交换机，把多达72个的线卡机箱连接起 来支持多达1152个40G的线卡。因此，CRS-1的多机箱互联的最大交换速率可达：1152×40Gbps×2= 92160Gbps=92Tbps。

下图所示为一个思科CRS-1 互联的实例图：

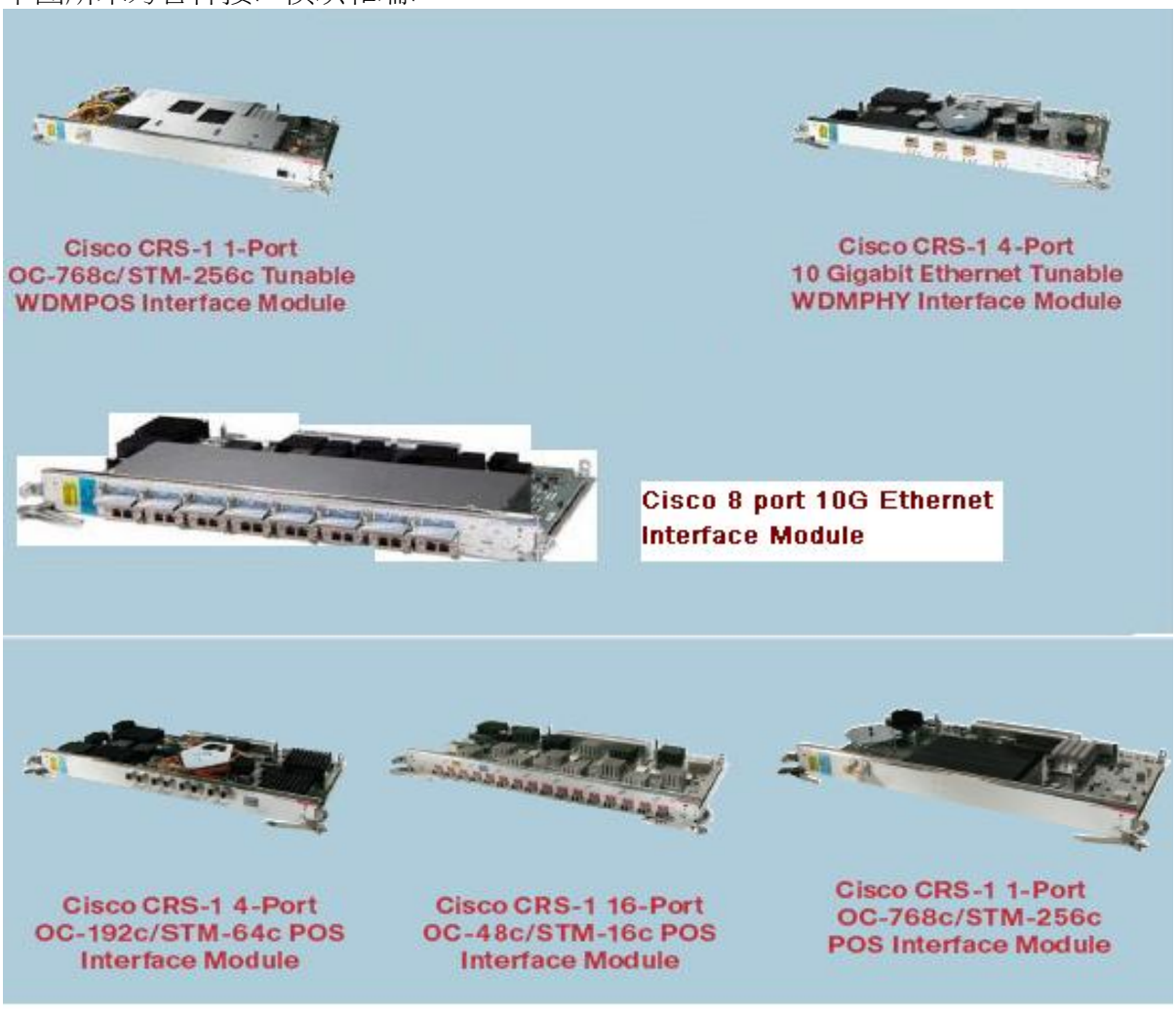


3. 端口配置

CRS-1产品家族的4，8和16线卡槽的线卡机（Line Card Shelf)可以单独作为路由器被安放在网络的节点上，也可以通过24槽的交换互联机（Fabric Shelf）互联形成一个强大的路由器阵列，提供超大规模的IP/MPLS报文交换能力。
从微观角度观察，其最基本的数据通信的粒度就是通过位于线卡机上的线卡（Line Card）的接口模块（Interface Module）所提供的。这里读者注意一个CRS-1的概念和术语：CRS-1上的线卡其实是包括3个部分（硬件）：接口模块卡（Interface Module），模块服务卡（Modular Services Card）和中间板（Midplane）。换言之，线卡是2个硬件卡通过插在中间板上组合形成。
CRS-1产品系列支持如下的接口模块：
×40Gbps端口的OC-768c/STM-256PoS卡。这个（40Gbps）卡上就是只有一个（40Gbps）端口（Port）。
×10Gbps端口的OC-192c/STM-64c PoS/DPT卡。这个（40Gbps）卡上有4个端口。4×10Gbps=40Gbps。
×10Gbps端口的Ethernet以太卡。这个卡上有8个10Gbps以太端口。40Gbps交换。可以Oversubscribe只利用4个端口。

×2.5Gbps端口的OC-48c/STM-16c PoS/DPT卡。这个（40Gbps）卡上有16个端口。16×2.5Gbps=40Gbps。
×40Gbps端口的OC-768c/STM-256 tunable WDMPOS卡。这个（40Gbps）卡就是只有一个（40Gbps）端口（Port）。
×10Gbps端口的Ethernet 以太tunable WDMPHY卡。这个（40Gbps）卡上有4个端口。4×10Gbps=40Gbps。
对于4，8，16卡槽的CRS-1线卡机，针对上述的端口，其配置相应为：
×4, 8, 或16 个OC-768c/STM-256 PoS端口
×16, 32, 或64个 OC-192c/STM-64c PoS/DPI端口
×32, 64, 或128个 10 Gigabit Ethernet以太端口
× 64, 128, 或256个 OC-48c/STM-16c PoS/DPT端口
× 4, 8, 或 16个 OC-768c/STM-256 tunable WDMPOS 端口
×16, 32, 或64个 10 Gigabit Ethernet tunable WDMPHY端口
当CRS-1线卡机通过最多可达8个Fabric交换机互联起来时，线卡机的数量可达72个。因此，**在最大容量下，这72个线卡机可以都是16槽的线卡机**。因此，CRS-1多机组的最大交换容量为：
× 72个线卡机
×（最多）8个Fabric交换机互联
×交换能力：**72×(16*40Gbps)*2=92160Gbps=92Tbps (16*40Gbps*2是一个16卡槽的线卡机的双向交换能力）**
×支持40Gbps接口模块：1152 个40Gbps接口模块（72×16=1152个40Gbps接口模块）
× 1152 OC-768c/STM-256 PoS 个端口
× 4608 OC-192c/STM-64c PoS/DPT 端口
× 9216 10 Gigabit Ethernet 端口
× 18,432 OC-48c/STM-16c PoS/DPT 端口
× 1152 OC-768c/STM-256 tunable WDMPOS 端口
× 4608 10 Gigabit Ethernet tunable WDMPHY端口

下图所示为各种接口模块和端口：



4. 线卡机配置

在讨论了CRS-1的接口模块（Interface Module）之后，对CRS-1能支持的端口（Port）配置就有了一个基本的了解。本节介绍这些接口模块是如何配置在线卡机（Line Card Shelf）之上的。在具体介绍和举例之前，读者要抓住一个概念，一个线卡机上的端口模块槽可以插入不同的40Gbps的端口卡，例如上OC- 768，OC-192，或者10G以太卡或其他端口卡。可以是各种混合模式。对系统而言，只要是支持的端口卡就可以。对系统而言，就是一个40Gbps线 速的报文端口硬件板子。不管你是光纤还是铜线，和层2的数据格式，一旦走到Fabric互联网络入口，都一样；都被打散成为Cell，做高速Fabric Switching了。
LCC线卡机内部的板子都是通过中间板（Midplane）互联的，不是Backplane（背板）。另外，对理解机箱板子布局， 抓住几个要点：
×端口卡（Interface Module）一定都是在机箱的正前面。
×路由处理器卡（Routing Processor Card）一定都是在机箱的正前面。
×交换卡 （Fabric Card）一定都是在机箱的后（背）面。
×模块服务卡（MSC）除了4槽的CRS-1线卡机，都是在机箱的后面。
×线卡机都是基于中间板（midplane）的设计，而非背板（Backplane）。

4槽线卡机
机箱的前面（Front），总共有10个卡槽。其中：
×4个40Gbps的端口卡槽（或叫做接口模块）
×4个支持40Gbps的模块服务卡（Module Service Card）槽。其中，每个MSC上有两个SPP网络处理器。关于SPP会在后续的章节里介绍。
×2个路由处理器卡（Routing Processor Card）槽 。在思科的网络产品里，路由处理器就是系统的控制平面。在Juniper的产品里，相应的概念是路由引擎（Routing Engine）。名称不同，表达的概念和功能是类似的--系统的控制平面。
上述的10个卡都通过中间板（Midplane），插入并连接在系统中。
×4个Fabric卡，在机箱的后面（Rear），插入中间板。从而与其他在 前面的10个板子连接，形成一个路由器系统。

下面是一个满载的4槽线卡机的例子：

思科4线卡槽机箱CRS-1路由器

8槽线卡机

机箱的前面（Front），总共有10个卡槽。其中：
×8个40Gbps的端口卡槽（或叫做接口模块）
×2个路由处理器卡（Routing Processor Card）槽 。上述的10个卡都通过中间板（Midplane），插入并连接在系统中。
×8个与端口卡对应的MSC模块服务卡和4个Fabric卡，在机箱的后面（Rear），插入中间板。从而与其他在 前面的10个板子连接，形成一个路由器系统。

下面是一个满载的8槽线卡机的例子：

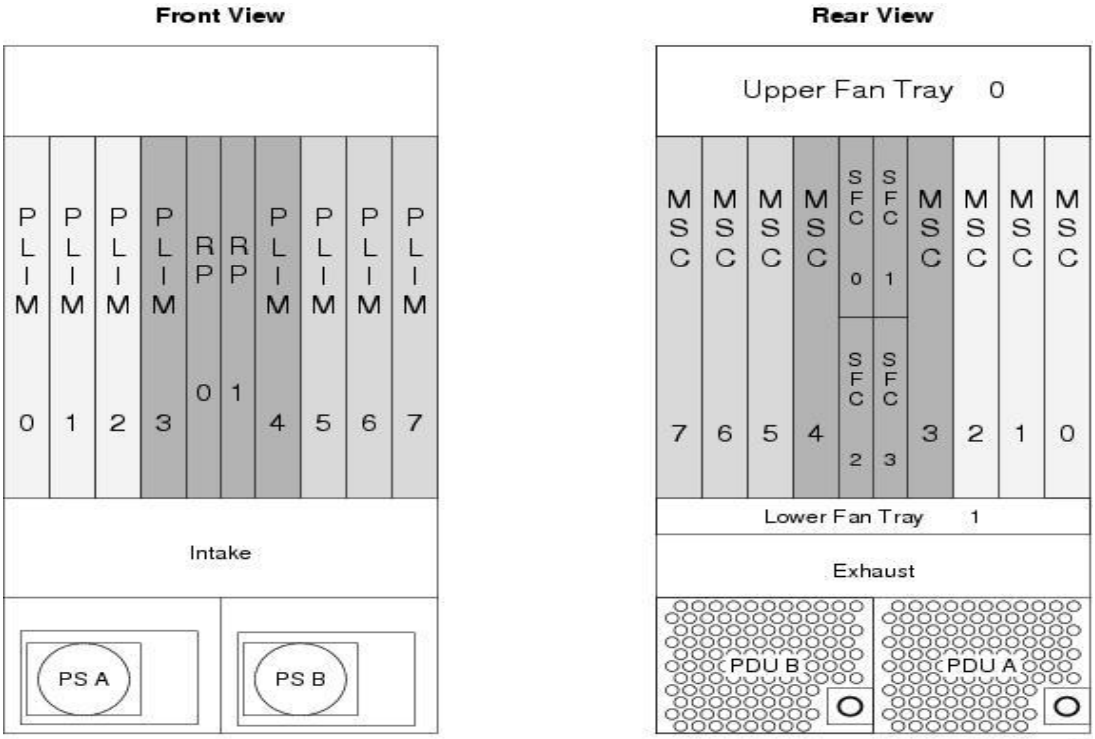
CISCO CRS-1 8-SLOT SINGLE-SHELF

8个40Gbps的端口卡槽（卡槽编号为1，2，3，4，7，8，9，10）。卡槽5，6是双控制平面RP（路由处理器卡）。

相应的8个支持端口卡的模块服务卡（MSC）是位于机箱的后面，通过并插入中间背板，与前面的10个板子互联成为一个系统。

这个8槽线卡机，1，2是两个OC-768卡。3，4是两个8端口的10G以太网卡。7，8，9是4端口的OC-192卡。10是16个端口的OC-48的卡。

思科8线卡槽CRS-1路由器



Load Zone 1

Load Zone 2

Load Zone 3

思科CRS-1 8槽线卡机前面和后面平面图

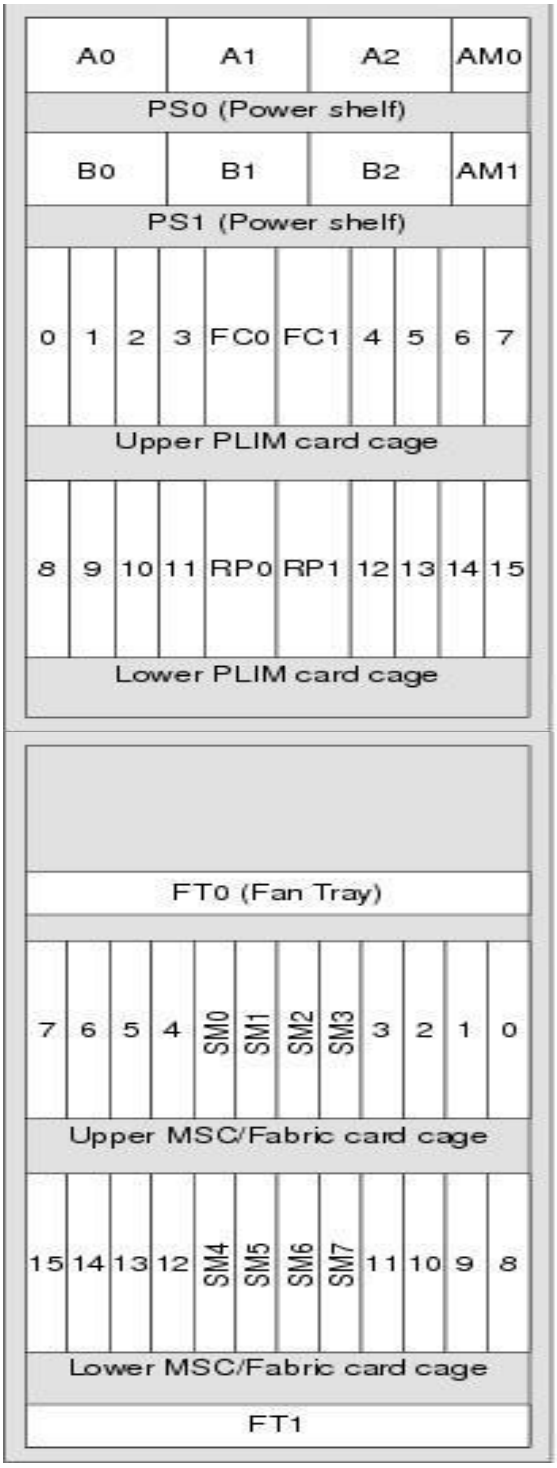
PLIM:就是端口卡，或者接口模块
RP：控制平面RP卡
MSC：配有两个SPP的模块服务卡。
SFC：Switch Fabric卡

16槽线卡机

机箱的前面（Front），总共有20个卡槽。其中：
上半部分：
×8个40Gbps的端口卡槽（或叫做接口模块）
×2个机箱控制卡
下半部分：
×8个40Gbps的端口卡槽（或叫做接口模块）
×2个控制平面的RP路由处理器卡。
×16个与端口卡对应的MSC卡和8个Fabric卡，在机箱的后面（Rear），插入中间板。从而与其他在 前面的20个板子连接，形成一个路由器系统。
下面是一个满载的16槽线卡机的例子：

思科16线卡槽CRS-1路由器

下面是CRS-1 16槽线卡机的前面和背面的平面示意图：



思科CRS-1 16槽线卡机前面（正面）示意图。
其中A0,A1,A2,AM0是电源模块。B0,B2,B2,AM1是另一组电源模块。

其他卡槽为16个PLIM（端口模块卡）

FC0和FC1为机箱风扇控制卡（Fan Controller）。

RP0和RP1为控制平面—路由处理器卡。

思科CRS-1 16槽线卡机的后面（背面）示意图。

共有16个MSC模块服务卡。

8个Fabric交换卡。

SM0-SM7.

对于CRS-1线卡机产品系列，所有的Fabric交换卡都是在机箱的后面。
1.通过内部的中间板连接内部系统的其他模块板卡。
2.通过光纤Cable进行CRS-1机组互联。（因此，当然不能把大把的光纤线缆暴露在机箱的正面。否则很难看）

5. 交换机配置

CRS-1产品系列中一个重要的组成部分就是其CRS-1的交换机FCC--Fabric Card Shelf。FCC产品的目的是通过提供一个或者多达8个FCC的交换机组，从而可以把2个，多达72个CRS-1线卡机（Line Card Shelf，简称LCC）互联起来，形成8个并行的高速交换平面（Switching Plane）。从理论上，CRS-1 多机箱技术（Multi-Shelf）可以提供高达92Tbps的报文交换容量。这种互联结构在工程实现上，机箱管理（Chasis Management)，系统软件上都很复杂。在目前为止，思科只能支持16槽的LCC，通过1，2或者4个FCC的互联拓扑。换言之，读者可以认为 92Tbps交换能力的互联，以CRS-1目前的体系结构而言其实是一个单纯的理论值。如果真的是72个LCC通过8个FCC互联，系统能否支撑（Sustain），笔者抱着谨慎乐观的态度。目前没有任何消息来源能证明思科自己内部是否做过这样的互联。其工程量非常大。能把所有的光纤正确的接对就 相当优秀，更不需要说IOS-XR的各种板卡的初始化，数据管理等更复杂的问题。

FCC本身是一个满载可以配备24个交换卡的机箱。在FCC的术语里面，这个交换卡称为S2卡。在以后的CRS-1互联的章节里，笔者会介绍S2卡的更多细节。从交换机配置的角度，FCC的配置是：

×正面24个S2交换卡。40Gbps的进出。上下两层。每层12个卡槽。

×正面2个机箱管理控制卡（Shelf Controller Gigabit Ethernet Card，简称SCGE）。

×后面24个OIM(Optical Interface Module)卡。

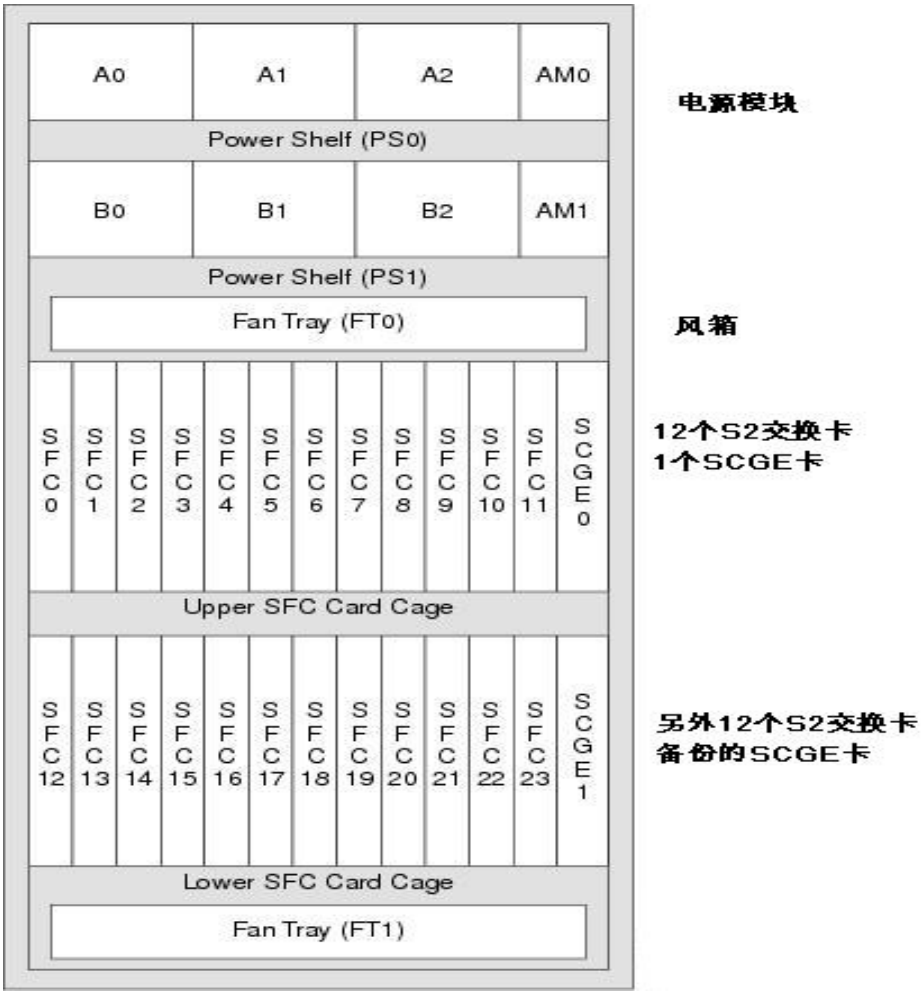
×后面2个OIM监视卡（OIM-LED）。

其中，后面的OIM卡就是用来互联LCC线卡机上的交换卡的。换言之，CRS-1机组互联是通过LCC交换卡上的适配器，然后通过光纤，连接到 FCC的OIM卡上的适配器或者接口上，而形成互联网络的。在LCC与FCC互联时，例如16槽LCC，其后面的交换卡是一种叫做S13的交换卡。而 LCC单独作为路由器时，其交换卡是S123交换卡。这是两个不同的物理卡。读者目前只需要了解在互联时，LCC要把S123的交换卡变成S13的卡就可 以了。

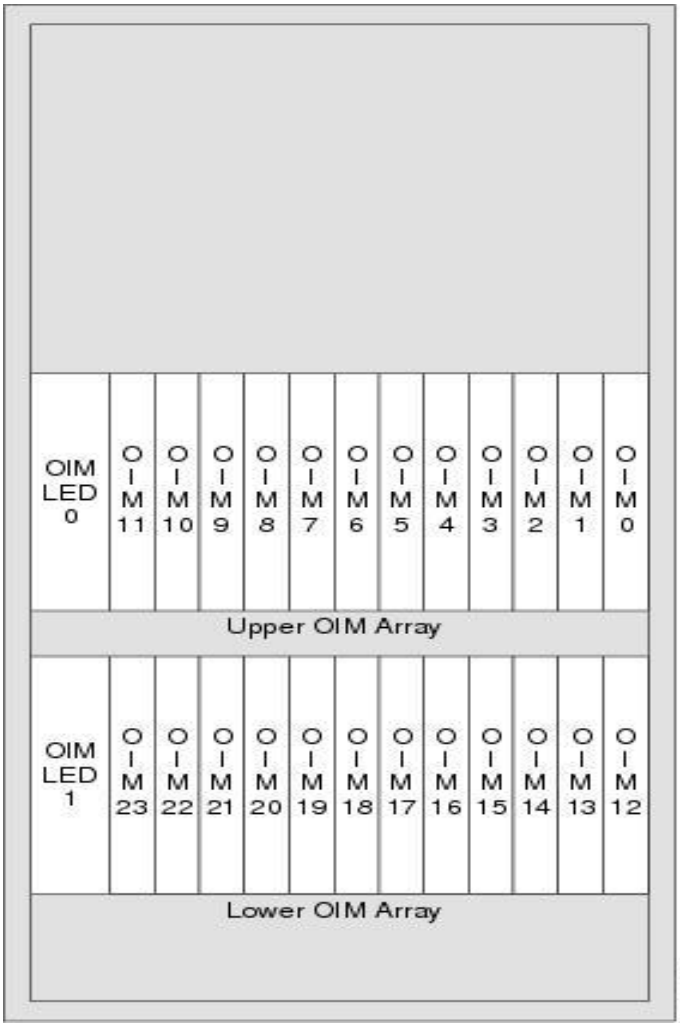
机箱管理控制卡SCGE是CRS-1系统，例如通过LCC上的控制平面RP来启动，管理，监测FCC的唯一途径。SCGE卡分为2种类型--2个GigE端口的SCGE和22个端口的SCGE。这里最大的区别不是端口数目，而是层2的switching功能。

如果一个系统配置的是2个端口的SCGE，系统需要另外配备思科的Catalyst switch。LCC的RP通过Catalyst交换机解决SCGE，从而达到控制FCC机的目的。而如果是最新的22个端口的SCGE，就不需要另外配 备Switch了。LCC的RP可以通过GigE以太端口直接连接SCGE。

下面是FCC的正面示意图：



下面是FCC的背面示意图：



OIM卡，用来与LCC的S13交换卡互联，形成CRS-1路由器阵列

一共24个OIM卡。每个卡上有9个OIM接口（适配器），用于光纤接入。

OIM LED卡用来监测OIM卡的状态。

6. 交换矩阵与交换平面

CRS-1最重要的部分可以说就是其交换矩阵和拓扑。交换矩阵把CRS-1的各个组成部分有机的结合在一起，形成一个高速的路由器。CRS-1的交 换是CRS-1体系结构里略微复杂和难理解的部分，下面分别对LCC和FCC的交换矩阵和拓扑做一些概念上的解释和阐述。希望读者能通过阅读相应章节之后 能从整体上对CRS-1的交换空间有一个清晰的把握。

首先，读者要了解CRS-1交换矩阵（Switch Fabric），交换平面（Switch Plane），交换卡（Switch Fabric Card）和之间的关系。本节介绍交换矩阵和交换平面，这两个逻辑实体。交换卡属于物理实体（实现），将在下一节更详细的介绍。

× **CRS-1交换矩阵（Switch Fabric）**： 交换矩阵由8个或者4个交换平面所组成。其中，16槽和8槽LCC：8个交换平面；4槽LCC：4个交换平面。交换平面之间是并行，没有依赖关系的。换言之，这些交换平面的关系是正交的。

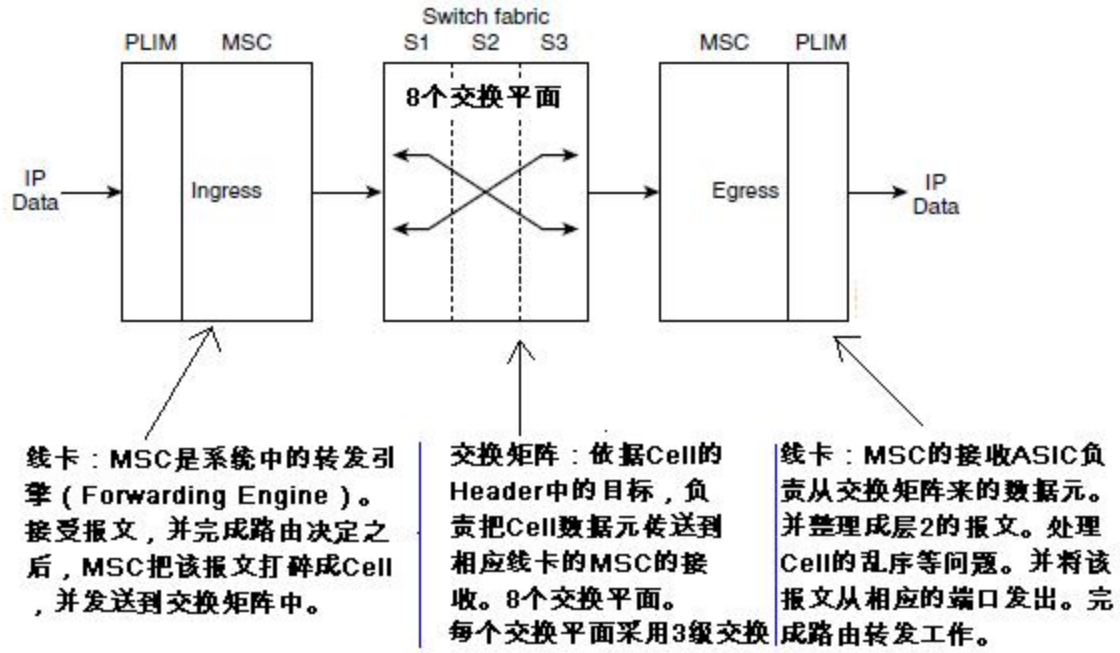
交换矩阵中通信的粒度是大小固定的数据元（Cell）。一个数据报文(Packet)在经过交换矩阵之前，线卡的Ingress MSC卡会把该报文分解为多个数据元。MSC会决定一个数据元被发送到哪个交换平面，并通过这个交换平面到达目标Engress MSC（这个MSC会汇总属于该报文的所有Cell，并恢复成一个Packet, 并通过端口卡（PLIM）上相应的端口（Port）把报文转发出去。一定要注意的是：属于同一个报文的不同的数据元可以（其实是通常）是通过不同的交换平 面抵达目的线卡的。但一个特定的Cell会而且只会通过一个选定的交换平面。在CRS-1，Ingress线卡上的MSC通常是通过轮询 （RoundRobin)的算法来决定使用哪个交换平面的。这里面的原因很简单直接，可以使得这8个或4个交换平面的负载均匀化，而避免某一个交换平面产 生阻塞丢掉数据元的情况。

× **CRS-1交换平面（Switch Plane）**：交换平面是交换矩阵的最小单位。一个交换平面能够实现了一个完整报文交换的功能。其中，每个交换平面采用3级（Stage）交换的体系结构。每个交换平面是通过在LCC机箱后面的交换卡来物理实现的。换言之，每个交换卡就是一个交换平面。

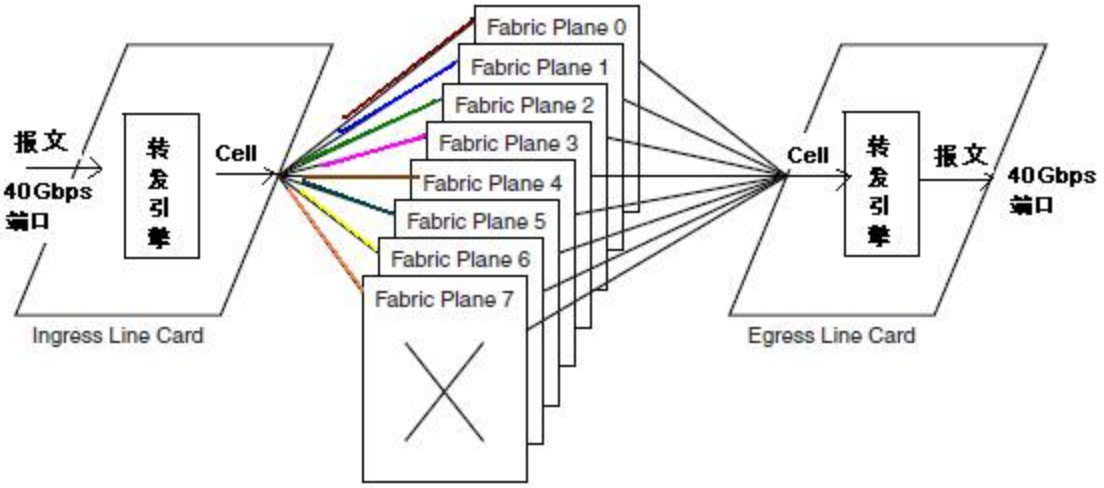
CRS-1系统中，理论上只需要一个交换平面就可以实现CRS-1路由器的转发交换功能。在CRS-1的实际运营上，系统要求至少两个交换平面来保 证系统路由转发交换。当然，这时系统已经不是线速的了。换言之，交换的能力不能支持众多40Gbps的线卡了。CRS-1的交换平面的交换率可以使得7个 交换平面就可以保证CRS-1系统的线速。

读者阅读到这里的时候，一定要深刻的认识如下的断言：**CRS-1系统中的任何（Each）一个线卡，通过中间板（midplane），可以连接到任何（Each）一个交换平面上。换言之，每个交换平面上，有来自所有线卡的输入。 每个交换平面上，都有通向任何一个线卡的输出。**

下图是CRS-1交换矩阵和交换平面的逻辑示意图。读者要注意的是，在谈论交换矩阵和交换平面的时候，都是一些逻辑概念和（或）实体。与具体的物理实现是无关的。或者说，可以通过许多技术方式来实现上述的交换矩阵和其组成部分交换平面。



CRS-1的交换矩阵示意图

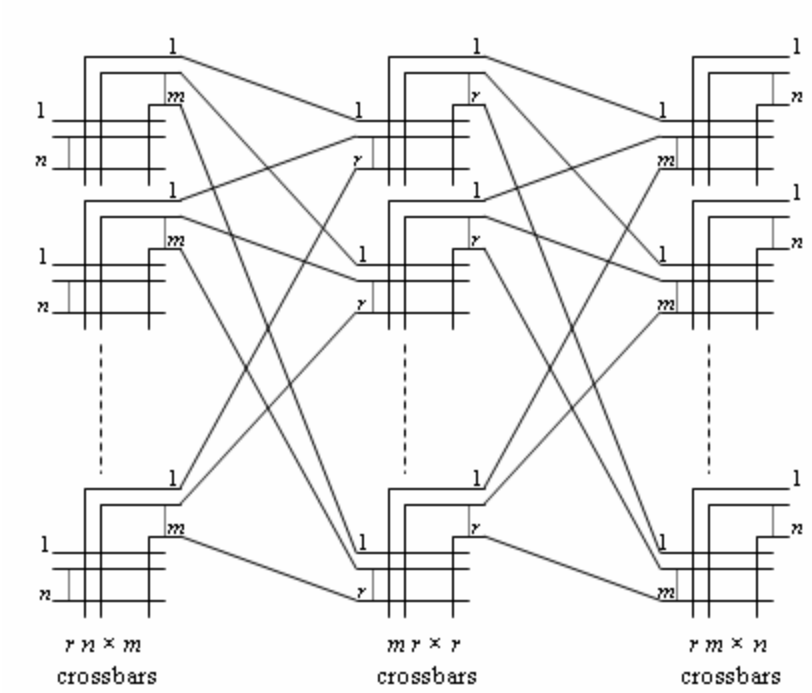


CRS-1交换矩阵有8个交换平面组成。同一个报文的Cell数据元可以通过不同的交换平面被传送到同一个目的线卡上。线卡负责整理Cells并形成完整的报文，并通过端口发出，离开CRS-1. 上图中的Ingress线卡与Engress线卡可以是同一个物理线卡。转发引擎（MSC）决定一切。

7. 物理交换卡（1）

CRS-1交换矩阵（Switch Fabric），交换平面（Switch Plane），交换卡（Switch Fabric Card)之间的关系是一个从属关系。前面两个是逻辑实体。交换卡属于物理实体（实现）。

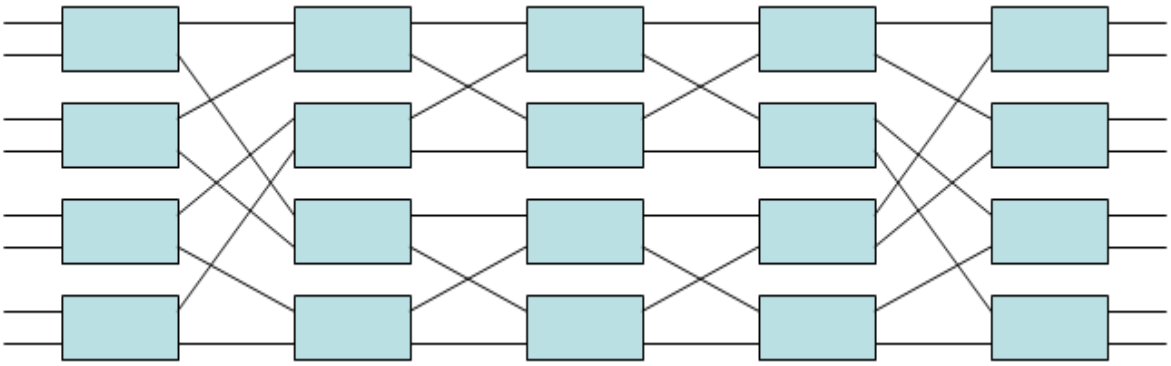
交换矩阵通过多个交换平面所组成。每个交换平面通过物理交换卡来实现。对于不同的CRS-1的LCC路由器，具体的数据关系如下：
16槽LCC：
交换矩阵：：==8个交换平面
1个交换平面：：==1个物理交换卡
【笔者注：16槽LCC配置8个交换卡槽。一个交换卡提供一个交换平面】
8槽LCC：
交换矩阵：：==8个交换平面
1个物理交换卡：：=2个交换平面
【笔者注：8槽LCC可配置4个交换卡槽。换言之，一个交换卡提供两个交换平面】
4槽LCC：
交换矩阵：：==4个交换平面
1个物理交换卡：：=1个交换平面
【笔者注：4槽LCC可配置4个交换卡槽。一个交换卡提供一个交换平面】
从上述数据读者可以得知。一个逻辑上的交换平面是通过一个（16LCC和4LCC）或者半个物理交换卡（8LCC）来实现的。
对于不同的LCC上的物理交换卡，其名称分别为：
16LCC：S123卡
8LCC： HS123卡
4LCC：QS123卡
CRS-1的交换不是一个全Cross-bar的结构，而是一个3级（Stage）Benes Network交换结构。关于Benes Network或者其更一般的交换结构拓扑Clos Network。简单而言，Close Network是一个通过3个参数（r,n,m)定义的一个3级交换拓扑，其3级结构分别为：输入（Ingress Stage），中间（Middle Stage）和输出（Engress Stage）。下图所示为一个通用的Cros网络拓扑模型：



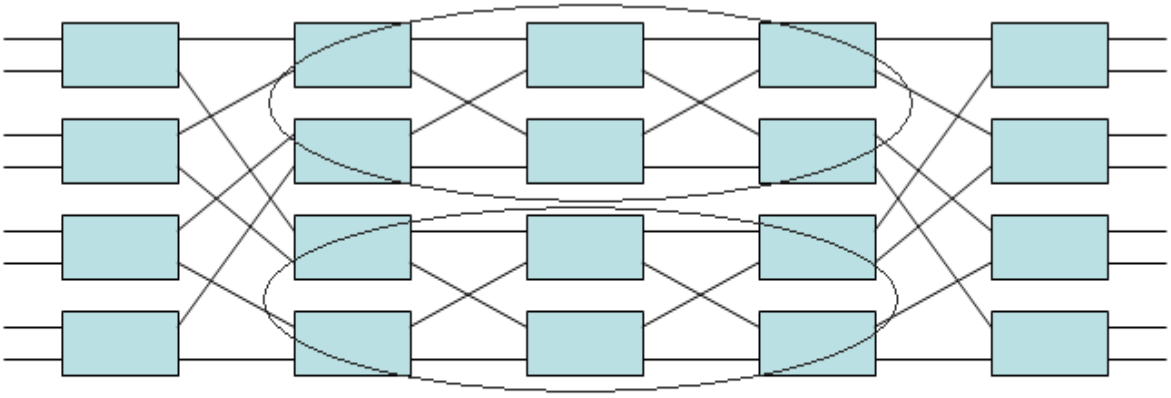
其中，r是输入（Ingress Stage）部件的个数。n是每个输入部件的输入接口数目。m是每个输入部件的输出接口数目。在理解Cros拓扑结构时，读者要抓住一个非常重要的概念：对于整个拓扑结构中的每个子部件，其是一个Cross-bar。例如，对输入而已，是有r个n×m个Crossbar。因为是r个输入部件，输出是m个， 因此，中间阶段一定是m个部件。为了形成cross-bar，很自然是r个输入，r个输出（除非做加速)。这也就是为什么中间阶段是m个r×r个 Cross-bar。在最后的输出阶段（Engress Stage），系统是一个逆转过程，通过r个m×n的cross-bar完成系统最后的交换过程。

关于Cros（Benes)拓扑的算法细节和量化分析，读者可以参阅相关文献。总而言之，Cros网络最大的优点是：相对一个没有中间交换过程的Cross-bar结构，对于要实现一个n×n的全交换，Cros网络所需要的连接节点数目要小的多。

Benes交换拓扑是Cros交换拓扑的一个特例，m=n=2。也就是说，在Benes交换拓扑中，每个Ingress和Engress子部件都是 一个 2×2的Cross-bar。读者要注意的是中间交换阶段（Middle Stage）是2个r×r的Crossbar。如果r是4，那就是2个4×4的crossbar。如果是16，那就是2个16×16的crossbar。 下图所示为一个8×8的Benes交换拓扑。



读者请注意，笔者说的2个4×4 的Crossbar是Cros拓扑的概念。在上述图中，就是中间3个Stage，共12个2×2的Crossbar。上半部分的两行（6个2×2的 crossbar）组成了一个4×4的Crossbar。下面两行（6个2×2的crossbar）组成了另外一个4×4的Crossbar。

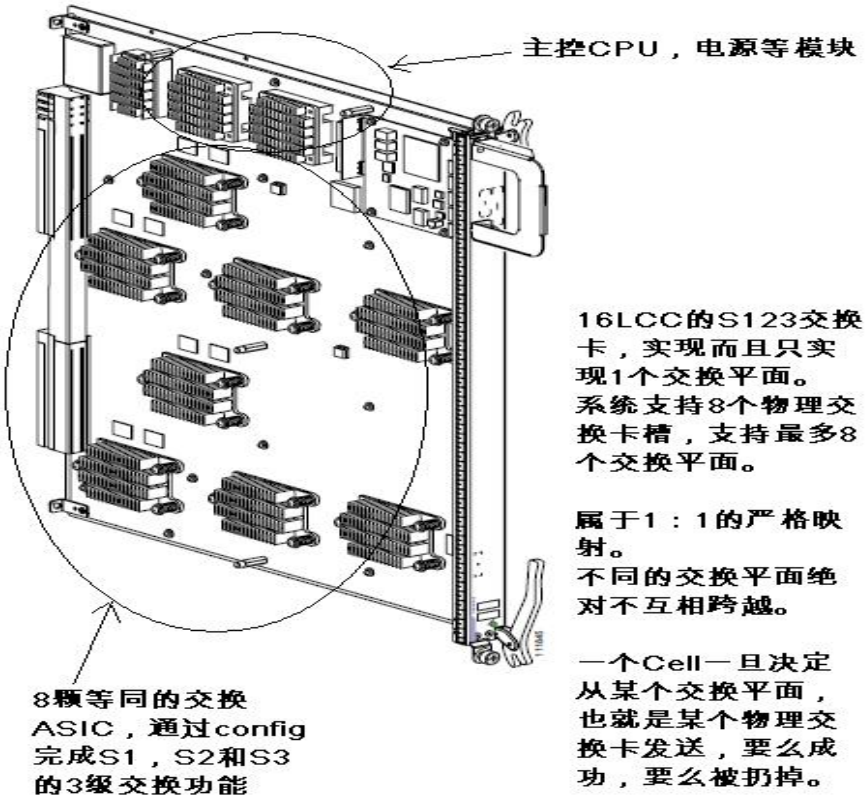


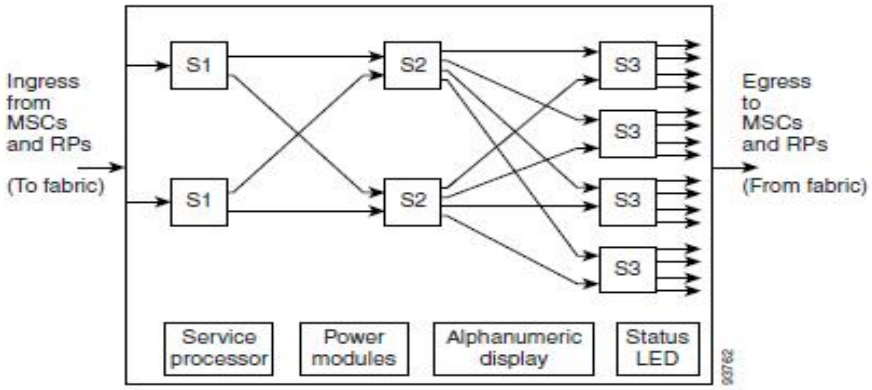
从3级结构Clos网络拓扑而言，中间是2个4×4的Crossbar。 m=n=2. r=4.

CRS-1的物理交换卡就是基于上述结构，而通过物理卡上的多个ASIC芯片而实现的。

8. 物理交换卡（2）

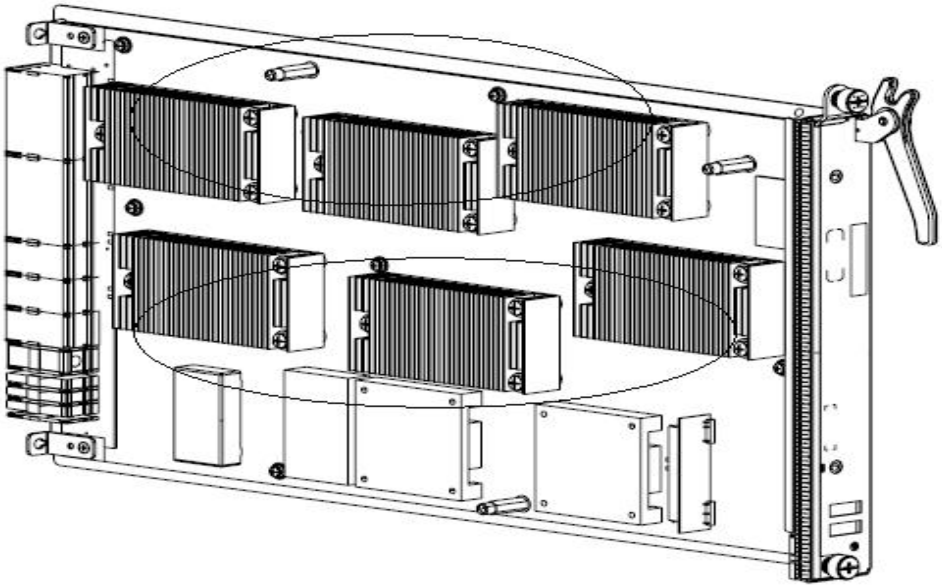
CRS-1 LCC路由器物理交换卡（Fabric Card）的概念模型是一个基于Benes Network拓扑的3级交换网络。在物理实现上，交换卡是通过卡上的多个ASIC芯片来完成3个阶段，基于固定大小Cell单元的数据交换。下面是一些基本的数据。
对于16槽的LCC，其交换卡（S123)上含有：2个S1，2个S2和4个S3的ASIC芯片，完成1个交换平面的功能。
对于8槽的LCC，其交换卡（HS123)上含有：2个S1，2个S2和2个S3的ASIC芯片，完成2个交换平面的功能。
对于4槽的LCC，其交换卡（QS123)上含有：1个S1，1个S2和1个S3的ASIC芯片，完成1个交换平面的功能。
下图所示分别为相应的物理交换卡:



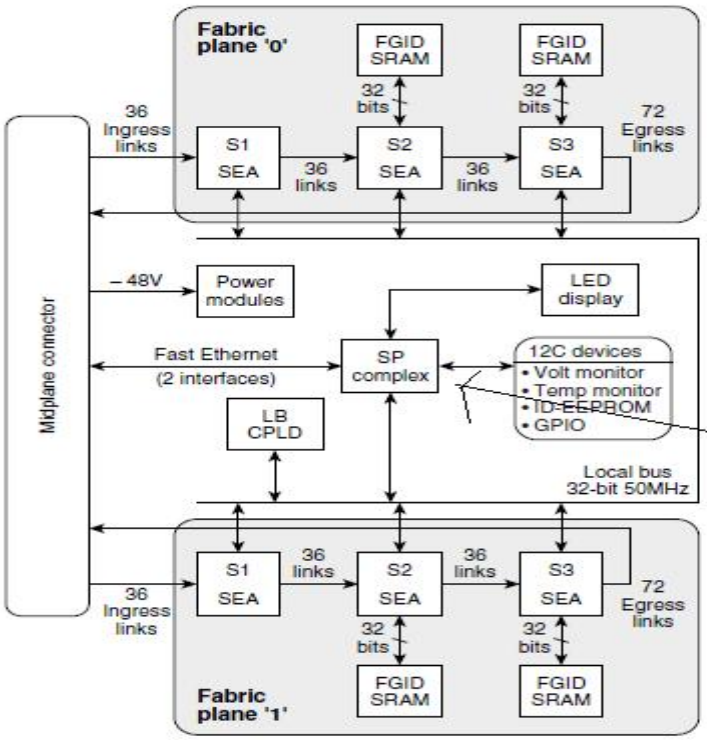


Each stage of the three-stage Benes switch fabric is implemented with the same switch element components. However, during system startup the components are programmed by Cisco IOS XR software to operate in S1, S2, or S3 mode, depending on their functions in the switch fabric. Each S123 switch fabric card contains two S1, two S2, and four S3 components.

注：16LCC物理交换卡芯片粒度概念图。对于8个交换芯片，都是一样的，并没有S1，S2和S3的区分。只是在操作系统初始化的时候才被选择担当不同的交换阶段的功能。



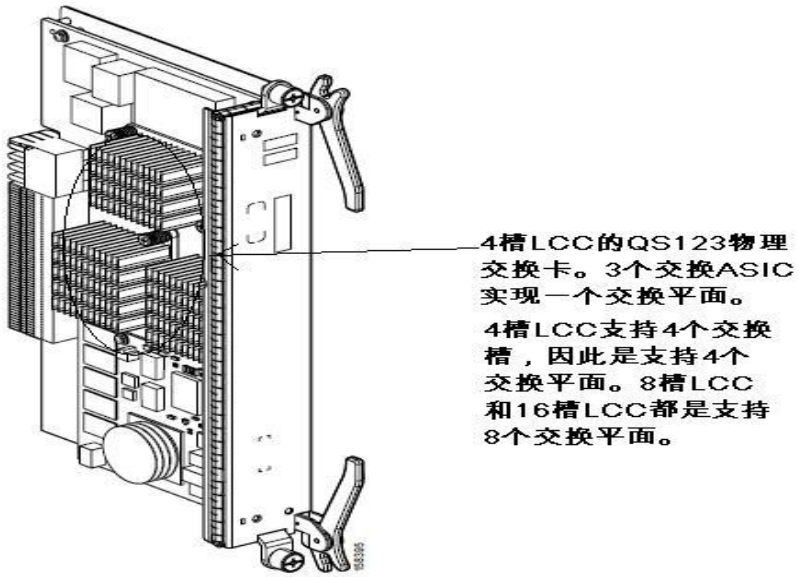
8槽LCC物理交换卡，6个交换ASIC，完成两个独立的，毫不相干的交换平面。每个交换平面有1个S1，1个S2和1个S3 ASIC芯片组成。



每个交换平面的S1 ASIC输入是36个2.5Gbps的Link。这一点非常重要。

板子上的主控CPU

8槽LCC的HS123物理交换卡，通过6个交换芯片，分为2组，实现两个独立的交换平面。



9. 物理交换卡（3）

在上述两节描述的物理交换卡（QS123，HS123和S123）是而且只是给CRS-1的独立路由器LCC的，例如，4槽，8槽或者16槽LCC。其实细心的读者可以从命名约定都可以得知，这3种交换卡的123后缀意味着在一个物理交换卡上，已经完成了交换平面的3级交换的各个阶段（Stage）。每个交换卡提供而且只提供系统8或者4个（对于4槽LCC而言）交换平面中的一个平面。

在CRS-1家族中，还有另外2个物理交换卡。其名称为S13卡和S2卡。这是对于CRS-1基于FCC的多机互联（LCC+FCC）的系统而专门使用的物理交换卡。

换言之，S13卡和S2卡是而且只是被用在基于FCC的多机互联系统中。请注意，笔者非常强调“基于FCC的”多机互联。其蕴含的信 息是，如果是单纯的多机LCC互联，例如双机16槽LCC的HA（Active/Passive）通过1G或者10G的数据端口互联，是不需要，也与 S13卡和S2卡毫不相干的。双机HA互联的结构，其实是一种容错结构。总体而言，是两个独立的路由器，只是在软件层面做NSR和ISSU等热备份工作。

CRS-1的FCC的多机互联的本质是：一个路由器！ 多个LCC加上一个或者多个FCC形成的一个系统是而且只是一个路由器。这个“只是一个路由器”既是逻辑上的理解，更重要的，也是物理上或者实际上的理解。

上述断言是对CRS-1系统非常重要的，希望读者一定要清楚的把握。一次性的把概念切入，从而避免混淆。

在一个路由器的前提下，理解下面要讲解的S13卡和S2卡就变得容易了。

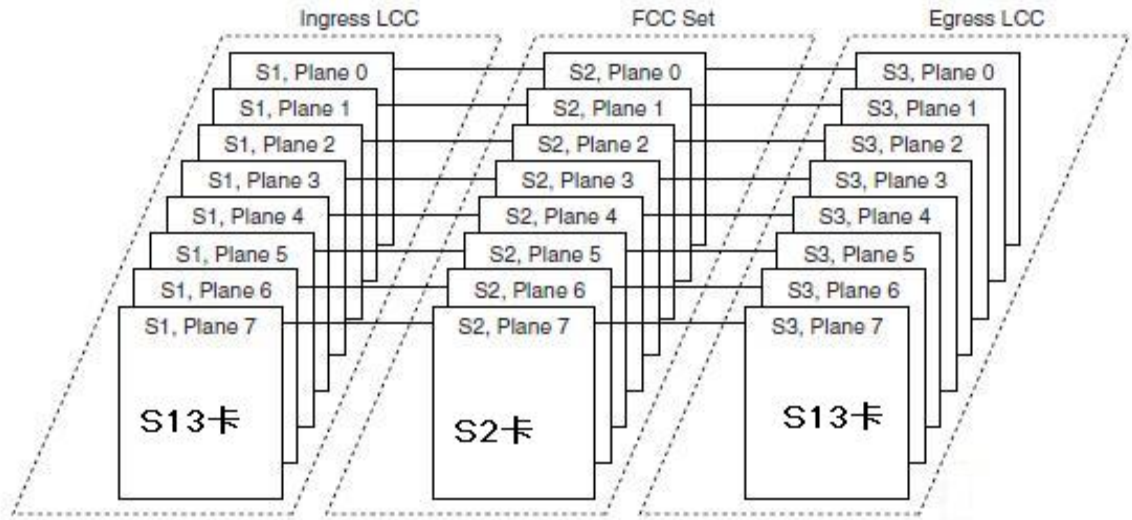
S13物理交换卡：

当16槽LCC用于FCC多机互联时，必须将单机时配置的S123卡拔出并替换掉。要插入S13交换卡。S13交换卡是CRS-1硬件系列中唯一能 与FCC通过光缆互联的设备。在S13物理交换卡上，一共有6个交换ASIC芯片 。其中2个完成S1阶段，4个完成S3阶段。读者比较S123卡，其实就 可以知道，其实就是去掉了两个S2阶段的ASIC交换芯片 。

在FCC多机互联的CRS-1系统中，LCC配置的是S13卡，而非S123卡。S13卡，顾名思义，具有3级交换的阶段1（Stage 1）和阶段3（Stage 3）的功能。
那么谁来完成基于3级交换的交换平面的阶段2的功能呢？
一个单独的物理交换卡--S2物理交换卡。
S2物理交换卡：
在S2卡上，含有6个交换ASIC芯片，也即S2芯片。S2卡是用而且只被用在24槽的FCC上。
到目前为止，基本上介绍了单机LCC系统下的交换矩阵，交换平面和相应的物理交换卡的概念，关系和具体的物理映射等。也介绍了在FCC多机互联下，CRS-1需要的另外两个特殊的物理交换卡--S13和S2卡。
笔者会在下节"FCC多机互联"中阐述：在FCC多机互联的体系结构中，如果延伸和映射交换矩阵，交换平面和物理交换卡这3个CRS-1重要的概念和含义 。 其实，忘却细节，读者如果能够理解笔者如下断言，就基本上把握住CRS-1的交换核心了：**LCC系统的交换是一个紧耦合系统（Tighly coupled System）。FCC多机系统 的交换是一个松耦合系统（Loosely coupled System）。**

10. FCC多机交换

FCC多机交换是LCC单机交换在逻辑上和物理上的无缝扩展。笔者用无缝（Seamless）这个词 想表达的意思就是，在CRS-1系统中，交换矩阵和交换平面，对这两个逻辑上的概念和实体，在LCC单机交换和FCC多机交换上，是一致的。唯一有区别的 是，LCC单机交换是各个线卡，交换卡都是插在并通过中间板（Midplane）进行连接的。而FCC多机交换，是把交换光缆（Switch Cable）通过LCC的S13交换卡的光缆接口与FCC的OIM（Optical Interface Module）光缆接口卡的光缆接口互联，从而组成交换矩阵和相应的8个交换平面。FCC机箱后面的OIM卡与FCC机箱正前方的S2物理交换卡是通过 FCC的中间板（Midplane）达到互联的。每个LCC的S13卡有3个光纤接口。每个FCC的OIM卡有9个光纤接口。一个FCC最大可以有24个 S2卡和相应的24个OIM卡。
CRS-1系统目前只支持16槽LCC的FCC互联。8槽LCC和4槽LCC不能支持FCC的互联。目前，CRS-1系统IOX最大配置容量可以支 持8个16槽LCC与1，2或者4个FCC互联。其中，2个或者4个FCC的目的是为了提高系统的容错性，例如可以通过配置把8个数据交换平面分布在2或 4个FCC上。但从理论和实践上而言，一个FCC的24个S2交换卡和相应的24个光缆接入卡OIM已经可以足够支持8个16槽LCC的FCC互联了。
下图是FCC互联的交换矩阵和交换平面逻辑图：



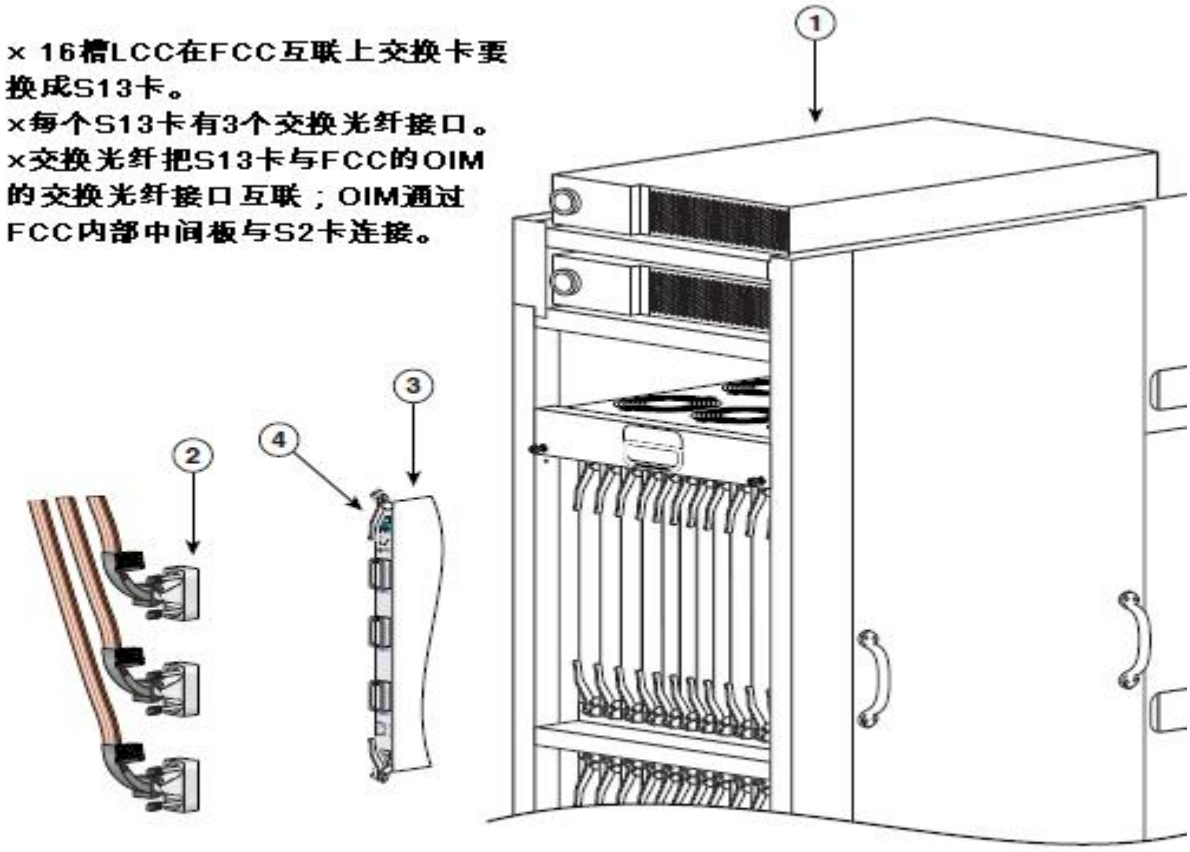
在FCC互联交换中，每个LCC的8个S13交换卡分别提供整个系统8个交换平面的S1和S3阶段。与LCC单机交换矩阵一样，一个S13卡属于而且只属于8个交换平面中的一个。任何一个数据单元（Cell）一旦交换，不会跨越交换平面。要么成功要么扔掉。一个数据单元通过S13卡的S1 ASIC芯片，然后通过光缆线，进入FCC的OIM卡，然后通过FCC的中间板，进入FCC的S2卡。S2卡再将该Cell转发给某个LCC的S13卡。从而被该S13卡上的S3 ASIC芯片介绍该数据单元。要注意的是，这个接收S13卡完全可以是当初该数据单元发送时在S1阶段的那个发送S13卡。道理就类似于Loopback一样。但是任何数据必须通过S2卡走一遍。不能是从一个S1 ASIC直接交换到同一个S13卡的S3 ASIC芯片上。

在上节中，我们介绍过S13卡和S2卡的物理配置情况。对于每个交换阶段，都是多个ASIC来负责承担。例如，6个ASIC在S2卡上。6个 ASIC在S13卡上（2个做S1；4个做S3）。同一个阶段的多个ASIC的目的和使用机制就是为了均匀负载，例如通过轮询使用算法。 而绝非是流水线的工作方式。

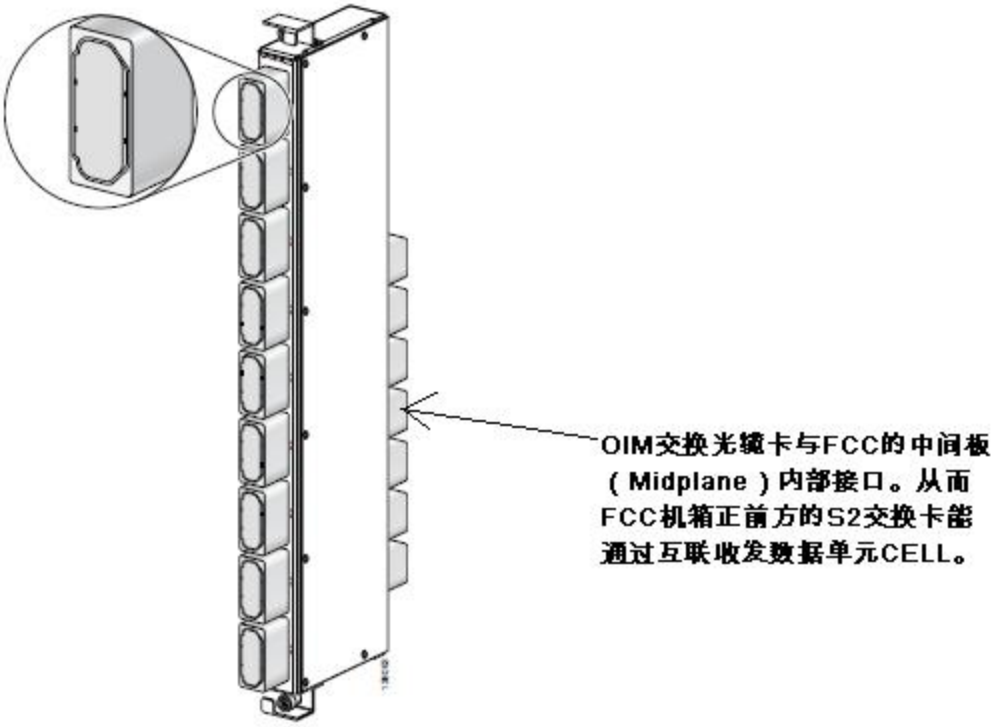
对于上述FCC交换矩阵和交换平面的理解中，有一点需要特别把握：

×一个交换平面可以是由有一个**S2**卡提供。也可以是通过**3**个**S2**卡来提供。换言之，笔者在上图中所标注的**“S2**卡**”**是一个逻辑实体。其可以是 一个**S2**物理卡，也可以是**3**个**S2**卡。目前为止，读者只要关心和认为这一点：每个交换平面，在**FCC**方面，是一个**S2**卡。关于为什么存在多个**S2**卡来提供 一个交换平面的原因和技术细节，会在将来的章节中解释。这些貌似复杂的变化其实都取决于**IOX**对系统的配置和现场多机互联的拓扑架构。

下面是LCC S13卡，FCC OIM卡光缆接口示意图：

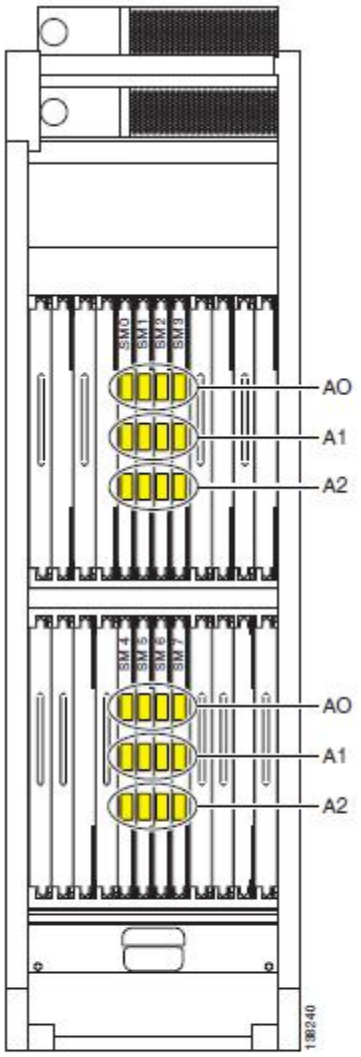


1	Line card chassis	3	S13 fabric card (FC/M)
2	Optical array cable connectors	4	S13 bulkhead array adapters

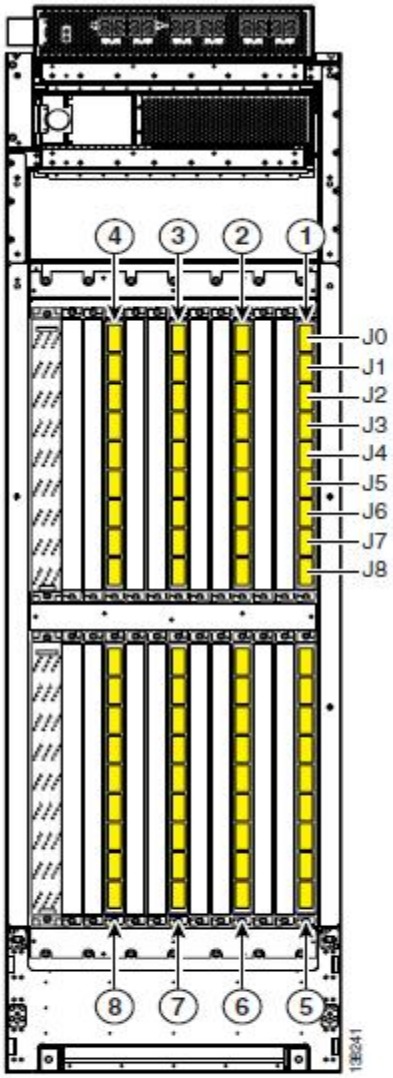


OIM交换光缆接入卡。提供9个交换光纤接口。

在一个S2交换卡负责一个交换平面的配置下，LCC S13卡的A0-2光缆依次接入在OIM的9个接口上。换言之，一个OIM卡可以接入3个S13卡的交换光缆。



16槽LCC，配有8个S13卡，用来FCC互联。每个13卡有3个交换光缆接口

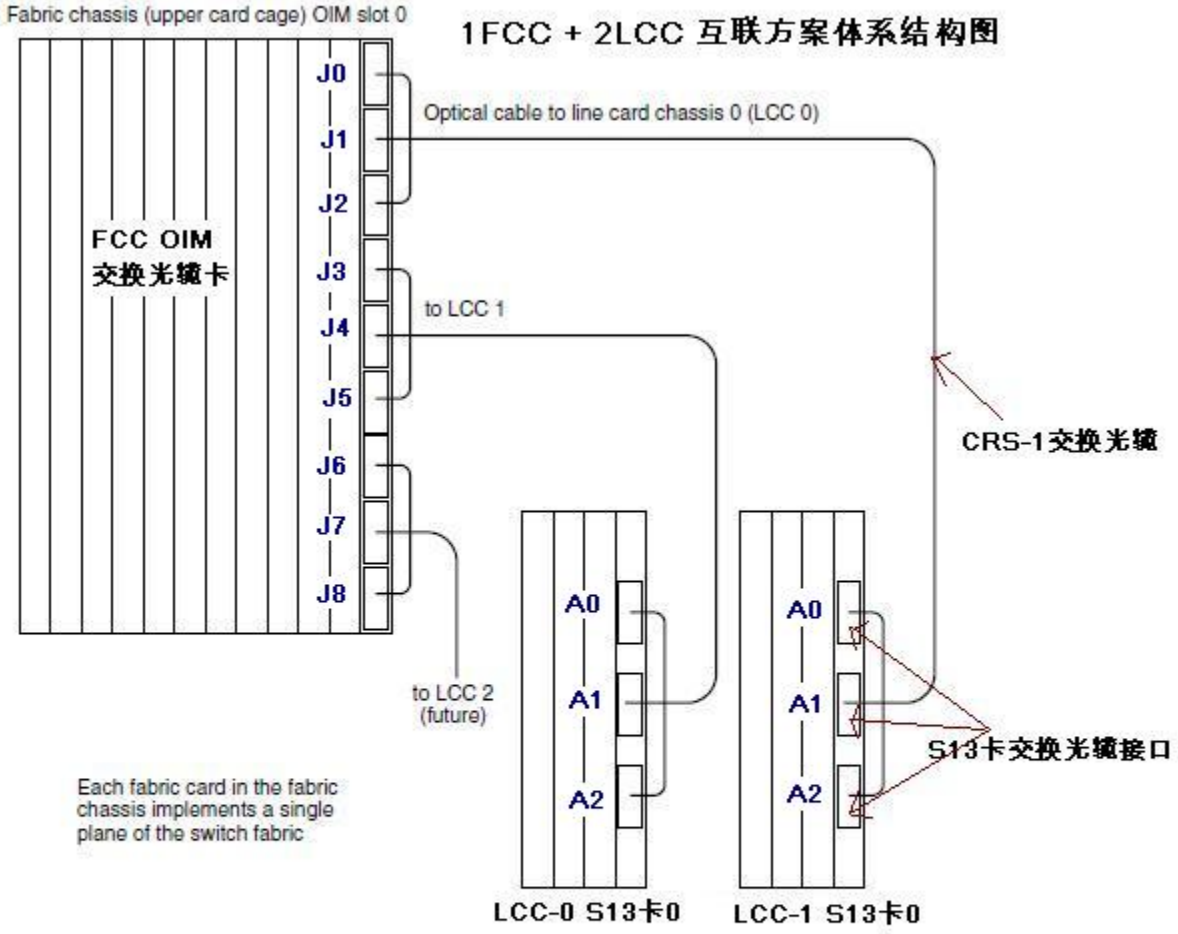


FCC。上图所示为配有8个S2交换卡和相应8个OIM卡。每个S2(OIM)提供一个交换平面。

在下面的章节中，笔者会力图非常清晰的把2LCC+1FCC, 4LCC+1FCC, 4LCC+4FCC等的拓扑解释给读者，一览CRS-1的FCC多机交换的奥秘。笔者会首先解释一个交换平面就是一个S2交换卡的情况。

11. FCC多机交换（2）

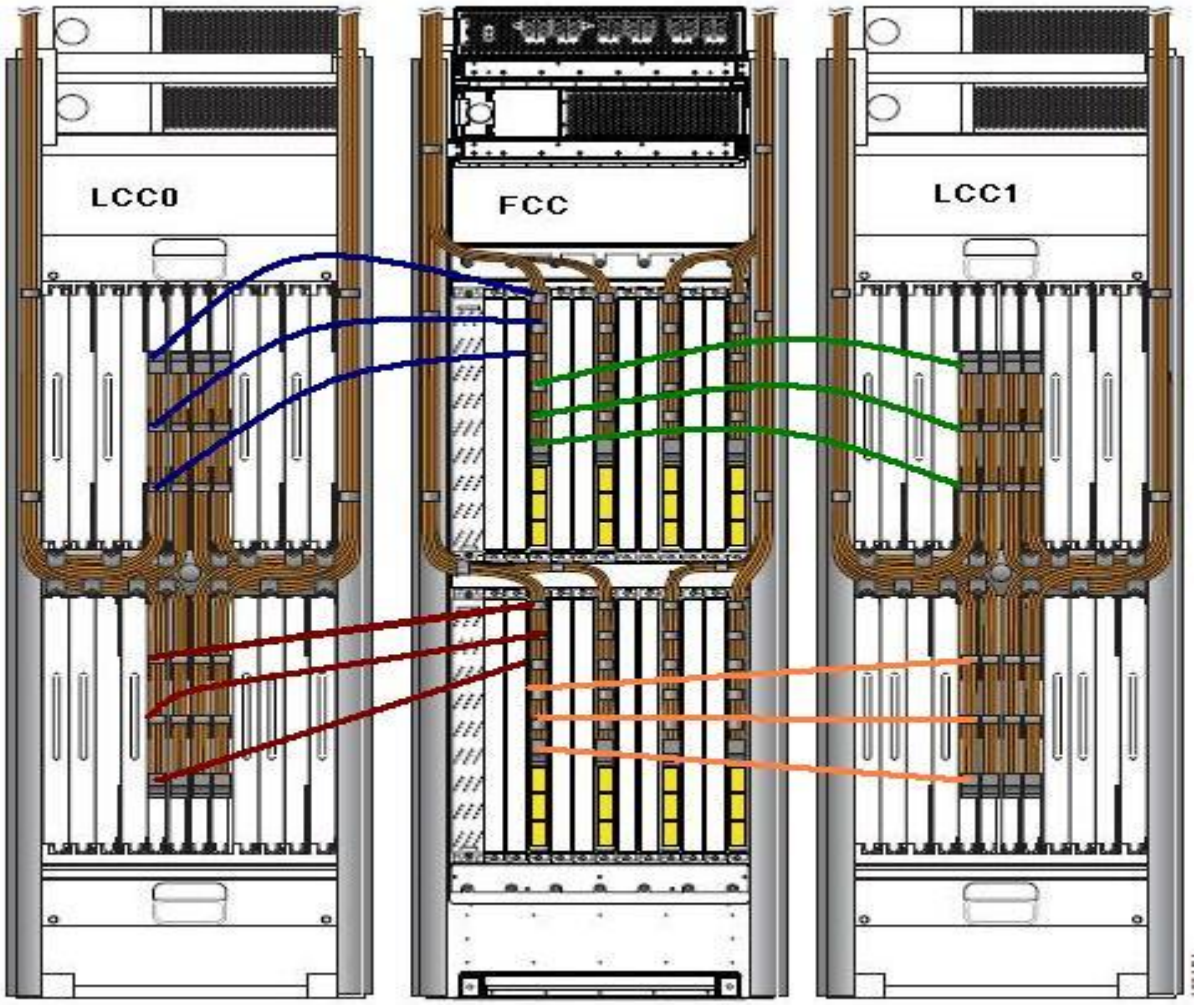
下图是1个FCC连接2个16槽的LCC的体系结构图。



LCC-0的S13卡0，LCC-1的S13卡0与FCC的S2卡0，通过FCC的OIM交换光缆卡和交换光缆组成系统中的一个交换平面。换言之，在这种配置下，每个S2卡完成一个交换平面。

对于上图的1FCC与2个LCC互联结构，一个OIM卡具备9个交换光缆接口。换言之，一个OIM卡可以支持3个S13卡的接入。原因很简单，每个 S13卡是3个交换光缆接口。一个OIM卡是与一个S2交换卡对应的。因此，一个S2交换卡可以把3个S13卡互联起来，组成一个交换平面。

下面是一个安装上述连接方案而形成的2LCC+1FCC互联的实例图。



CRS-1 2LCC+1FCC互联实例图。

在上面图中，系统中的16个S13交换卡（每个LCC有8个S13卡）通过交换光纤互联在FCC机箱后面的OIM卡的接头上。通过这样的物理互联，**这个CRS-1多机互联形成8个松耦合的独立的交换平面。每个平面有一个LCC0的S13卡，一个FCC的S2卡和一个LCC1的S13卡组成。当然，还包含其中的光纤线缆。**

在LCC机器中，任何一个线卡（MSC）可以通过中间平面（Midplane）连接到其本地LCC的8个S13卡的任何一个S13卡，数据可 以从线卡的MSC抵达任何一个S13卡。因此，在FCC互联的情况下，可以得知，任何一个数据单元从任何一个LCC的任何一个线卡，跨越互联光缆，抵达 FCC的S2交换卡，然后再抵达另外一个或者同一个LCC的一个S13卡的S3芯片，再抵达目的地线卡MSC。

图中FCC的OIM卡上黄色的部分是没有用的，空闲的3个互联光纤接头。

另外，图中标记的蓝，绿，红和橙色线段不是物理光缆，而是笔者标注的逻辑示意线段，用来表明S13卡，OIM卡之间的接头关系。真正的物理光缆是那 些褐色的，整整齐齐的物理光缆。例如，蓝色线表明LCC0的S13卡（编号：SM0）的光缆接头A0，A1和A2分别接在了FCC的OIM卡（编 号：OIM9）的J0，J1和J2上；绿色线表明LCC1的S13卡（编号：SM0）的光缆接头A0，A1和A2分别接在了FCC的OIM卡（编 号：OIM9）的J3，J4和J5上。红色线表明LCC0的S13卡（编号：SM7）的光缆接头A0，A1和A2分别接在了FCC的OIM卡（编 号：OIM21）的J0，J1和J2上；橙色线表 明LCC1的S13卡（编号：SM7）的光缆接头A0，A1和A2分别接在了FCC的OIM卡（编号：OIM21）的J3，J4和J5上。

读者可能会思考，FCC的OIM卡有24个。上下两排。对上图而言，上排，OIM卡/插槽用的是9，6，3，0（通过 J0，J1，J2，J3，J4，J5接口，依次链接两个LCC的0，1，2，3共2×4=8个S13交换卡。） 。下排用的是12，15，18，21（通过J0，J1，J2，J3，J4，J5接口，依次链接两个LCC的4，5，6，7共2×4=8个S13交换 卡。）。

这是为什么呢？

原因是容错。为了FCC最大程度的避免单点失败（Single Point of Failure）。

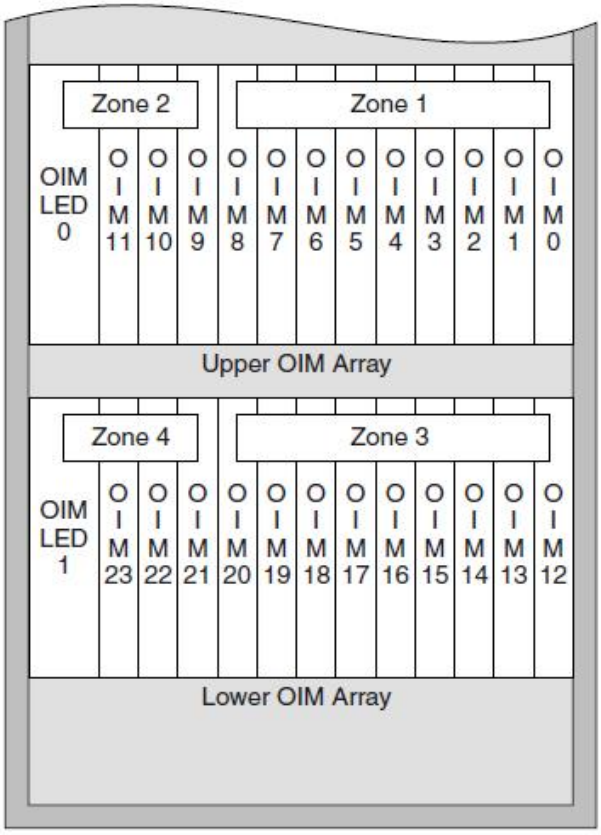
这里面涉及到FCC的电源系统。

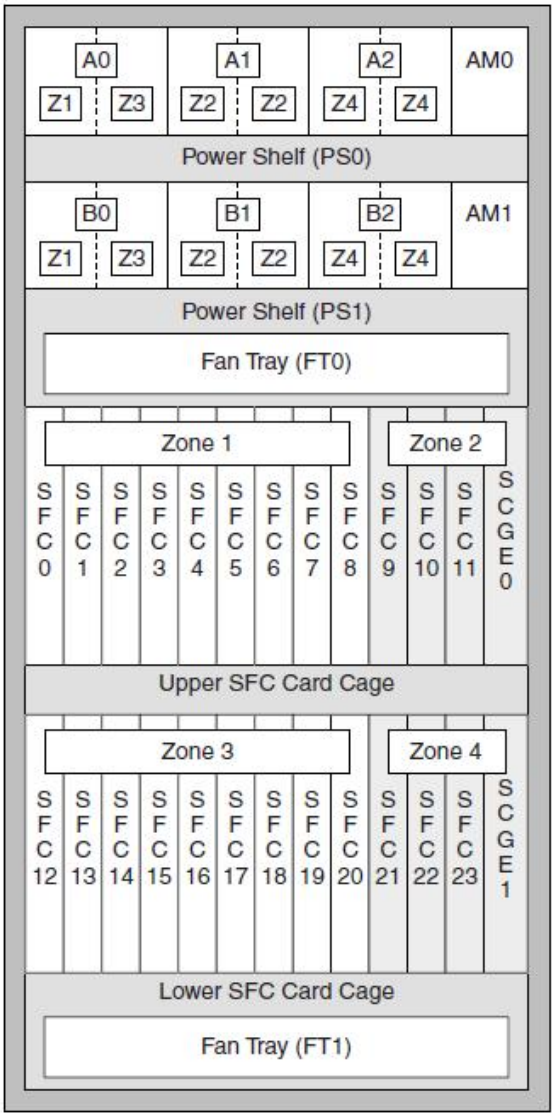
在FCC机箱中，电源系统是很复杂的一个部分。对于24个OIM卡而言，电源系统分为4个区域（Zone）。

对于电源而言，一个Zone的电源出错，就意味着该Zone内所有的OIM卡都失去了电源支持。

所以，作为系统架构设计，如果有可能，是要将OIM卡的使用分布在不同的电源区域，从而使得系统的容错性得到最大的体现。

下图所示为FCC对于OIM卡和S2交换卡的电源控制分布图：





FCC电源系统分区概念图。左图为FCC机箱背面OIM卡部分的电源4区域（zone）分布。OIM0—OIM8卡槽共享一个电源控制。OIM9-OIM11共享一个；OIM12-OIM20共享一个；最后，OIM21-OIM23共享另外一个电源控制。一旦某个电源控制失效，其管辖的所有OIM卡就断电不能工作。

右图为FCC机箱正面S2交换卡的电源控制图。同理，对于S2卡，分为4个电源控制区域。SFC0-SFC8一组；SFC9-SFC11一组；SFC12-SFC20一组。最后SFC21-SFC23共享一个电源控制。

12. FCC多机交换（3）

在上述两节描述的物理交换卡（QS123，HS123和S123）是而且只是给CRS-1的独立路由器LCC的，例如，4槽，8槽或者16槽LCC。其实细心的读者可以从命名约定都可以得知，这3种交换卡的123后缀意味着在一个物理交换卡上，已经完成了交换平面的3级交换的各个阶段（Stage）。每个交换卡提供而且只提供系统8或者4个（对于4槽LCC而言）交换平面中的一个平面。

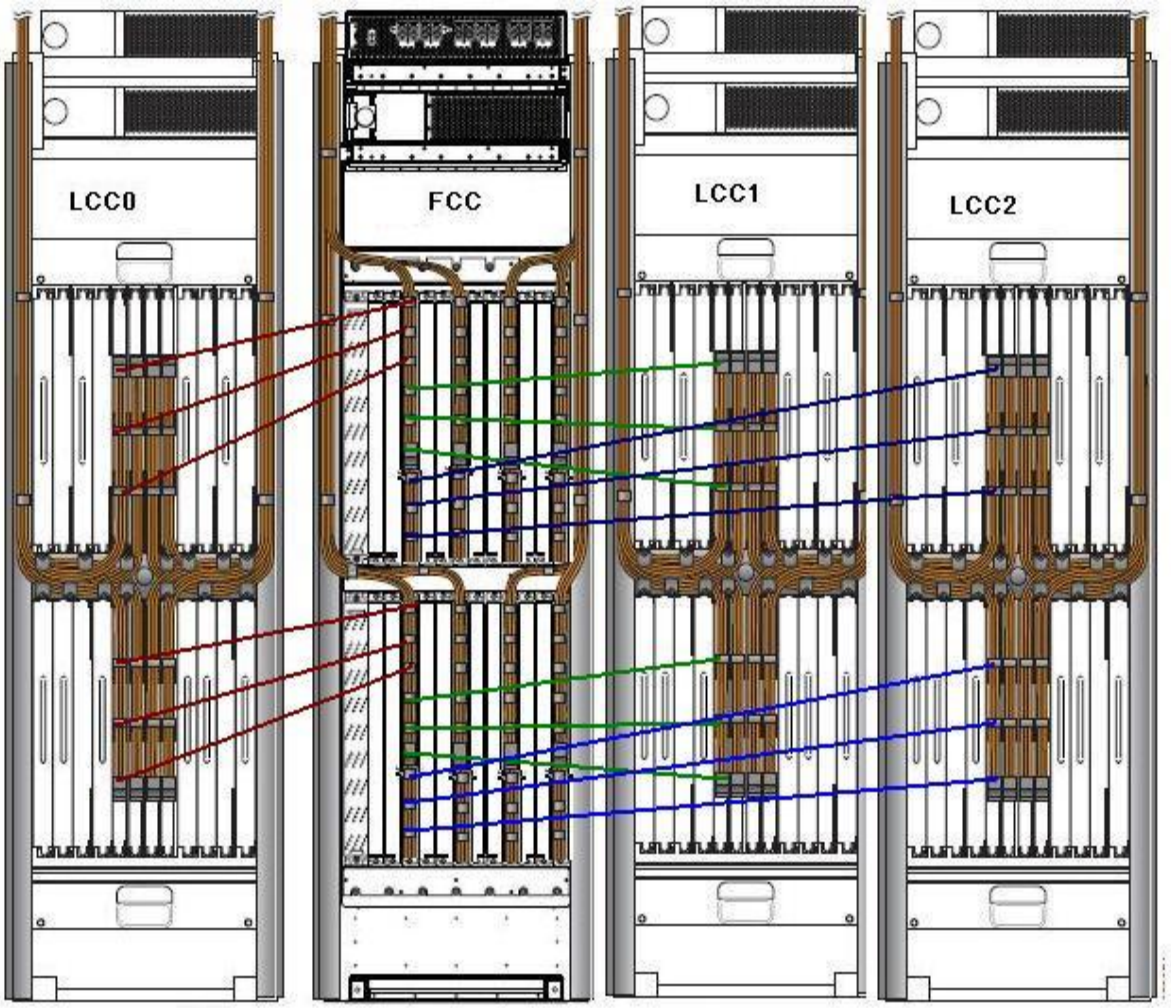
在CRS-1家族中，还有另外2个物理交换卡。其名称为S13卡和S2卡。这是对于CRS-1基于FCC的多机互联（LCC+FCC）的系统而专门使用的物理交换卡。
换言之，S13卡和S2卡是而且只是被用在基于FCC的多机互联系统中。请注意，笔者非常强调“基于FCC的”多机互联。其蕴含的信 息是，如果是单纯的多机LCC互联，例如双机16槽LCC的HA（Active/Passive）通过1G或者10G的数据端口互联，是不需要，也与 S13卡和S2卡毫不相干的。双机HA互联的结构，其实是一种容错结构。总体而言，是两个独立的路由器，只是在软件层面做NSR和ISSU等热备份工作。
CRS-1的FCC的多机互联的本质是：一个路由器！ 多个LCC加上一个或者多个FCC形成的一个系统是而且只是一个路由器。这个“只是一个路由器”既是逻辑上的理解，更重要的，也是物理上或者实际上的理解。
上述断言是对CRS-1系统非常重要的，希望读者一定要清楚的把握。一次性的把概念切入，从而避免混淆。
在一个路由器的前提下，理解下面要讲解的S13卡和S2卡就变得容易了。
S13物理交换卡：
当16槽LCC用于FCC多机互联时，必须将单机时配置的S123卡拔出并替换掉。要插入S13交换卡。S13交换卡是CRS-1硬件系列中唯一能 与FCC通过光缆互联的设备。在S13物理交换卡上，一共有6个交换ASIC芯片 。其中2个完成S1阶段，4个完成S3阶段。读者比较S123卡，其实就 可以知道，其实就是去掉了两个S2阶段的ASIC交换芯片 。
在FCC多机互联的CRS-1系统中，LCC配置的是S13卡，而非S123卡。S13卡，顾名思义，具有3级交换的阶段1（Stage 1）和阶段3（Stage 3）的功能。
那么谁来完成基于3级交换的交换平面的阶段2的功能呢？
一个单独的物理交换卡--S2物理交换卡。
S2物理交换卡：
在S2卡上，含有6个交换ASIC芯片 ，也即S2芯片 。S2卡是用而且只被用在24槽的FCC上。
到目前为止，基本上介绍了单机LCC系统下的交换矩阵，交换平面和相应的物理交换卡的概念，关系和具体的物理映射等。也介绍了在FCC多机互联下，CRS-1需要的另外两个特殊的物理交换卡--S13和S2卡。
笔者会在下节“FCC多机互联”中阐述：在FCC多机互联的体系结构中，如果延伸和映射交换矩阵，交换平面和物理交换卡这3个CRS-1重要的概念和含义 。 其实，忘却细节，读者如果能够理解笔者如下断言，就基本上把握住CRS-1的交换核心了：**LCC系统的交换是一个紧耦合系统（Tighly coupled System）。FCC多机系统 的交换是一个松耦合系统（Loosely coupled System）。**

13. FCC多机交换（4）

CRS-1系统，其理论最大满负载是72个LCC的CRS-1多机互联。
在关于CRS-1端口配置中，我们谈到，72个满负载的LCC互联的CRS-1系统，其端口数目惊人，可以支持1152 OC-768c/STM-256 PoS 个端口，4608 OC-192c/STM-64c PoS/DPT 端口，或9216 10 Gigabit Ethernet 端口，18,432 OC-48c/STM-16c PoS/DPT 端口，1152 OC-768c/STM-256 tunable WDMPOS 端口，或者4608 10 Gigabit Ethernet tunable WDMPHY端口。

在上述这样一个**72**台**LCC**互联形成的路由器中，作为一个路由器，其最基本的一个要求是：确保一个数据报文（**Packet**）从任何一个线卡的任何一个端口，能够通过路由路径寻找，从这**72**个**LCC**的任何一个线卡的某个端口被**CRS-1**转发（**Forwarding**）出去。
在理论上，为了达到上述目标，1个FCC，2个FCC，。。。，8个FCC的情况下，如何将这72台LCC机器的交换卡互联？
这里面最关键的还是要牢牢抓住**CRS-1**支持**8**个独立的交换平面。**8**个独立的交换平面是**CRS-1**的纲，**LCC**等是**CRS-1**的目。纲举目张。
首先来回顾一下当FCC的每个**S2**交换卡（或者说**OIM**卡）作为一个交换平面的情况。
在2LCC+1FCC的情况下，每个LCC全负载，也就是说，上架8个S13卡（请参阅16槽LCC的交换卡槽结构与配置）。为了实现8个平面互 联，FCC要贡献出8个S2（OIM）卡。每个卡出任一个交换平面。每个LCC的S13卡都通过其光纤接口A0-A2分别依次接到相应的OIM卡的光纤接 口J0-J2和J3-J5上。OIM卡的光纤接口J6-J8闲置不用。
在上述配置下，任何一个LCC的线卡都可以通过其S13卡<-->A0-A2接口-<-->OIM卡<-->S2卡<-->S3卡来抵达其他的一个LCC机器的线卡。
上述拓扑下，系统升级，要加一台LCC，形成3LCC+1FCC，如何互联？
很简单，把这台新的LCC的8个S13卡的A0-A2分别接到FCC的8个OIM卡的J6-J8上。
上述拓扑下，系统再升级，要再加一台**LCC**，形成**4LCC+1FCC**，如何互联？
读者的直觉有可能是，一个FCC支持24个S2交换卡（24个OIM卡）。在上述拓扑下，FCC只用掉了8个S2（OIM）卡，还有16个呢。FCC再贡献出来8个S2和8OIM卡，然后把这个新的第4个LCC 的S13卡的A0-A2接到这个新增加的OIM卡的J0-J2上即可。
这样的接入能使得系统达到“确保一个数据报文（**Packet**）从任何一个线卡的任何一个端口，能够通过路由路径寻找，从这**72**个**LCC**的任何一个线卡的某个端口被**CRS-1**转发（**Forwarding**）出去。”吗？

下图是笔者描绘的一个3LCC+1FCC的拓扑并通过这个实例来解释这种拓扑结构扩容的问题。



3LCC+1FCC互联拓扑(一个S2(OIM)卡作为一个交换平面)

在图中，笔者勾画出了3个LCC的两个交换平面的连接拓扑，分别是S13卡的SM0和SM4。有兴趣的读者可以依次把LCC剩下的7个S13卡通过FCC相应的OIM卡互联。我们会认识到，这种连接只能保证3个LCC的互联。

现在系统扩容，加一个LCC。自然，这个新的LCC也有8个S13卡和相应的光纤A0-A2接口。如何接入？

是的，FCC拥有24个OIM卡和相应的24个S2交换卡，而且目前其实只被用了8个，还剩下了16个OIM卡槽和16个S2卡。

但我们发现，即使你现在从FCC再贡献8个OIM卡，并且把在第4个LCC的8个S13卡的A0-A2都通过光纤接到相应的OIM卡的J0-J2 上，但问题出来了：系统中目前的3个LCC的S13卡的光纤接口A0-A2都已经被用光了！！1，2，3LCC的8个S13卡都已经被全部接到了上图中的 8个OIM卡上，并形成了8个交换平 面。FCC为了这个第4个新增加的OIM卡没有与任何来自其他LCC的S13卡的A0-A2互联。其他LCC的S13卡的A0-A2光纤接头都已经被全部 使用完了；相应的 OIM卡的9个J0-J8接头也被全部占用殆尽了！

因此，这样的接入使得这第4个LCC形成了一个孤岛。没有与任何其他的LCC构成联通。

在3LCC+2FCC，3LCC+4FCC的情况下升级为4LCC+2FCC和4LCC+4FCC，结果是一样的。单纯的从FCC多余的OIM卡拿 出一些来连接新的LCC，只会导致这个新的LCC是一个数据交换的孤岛。其他LCC的数据没有办法抵达（Reach）这个新LCC的S2卡和相应的S13卡。其他LCC的S13卡和相应的S2（OIM）卡都已经全满负载了（J0-J8都被全部用光了），从而导致这个新的LCC的S13卡无法加入相应的8个 交换平面。

不同交换平面 的数据是永远不会交叉转发的！这是CRS-1的一个原则（Rule)。

换言之，在一个S2卡作为一个交换平面的情况下，S2卡上的S2ASIC会而且只会

× 接受来自S13卡上的S1 ASIC芯片 转发的数据单元；

×转发数据单元去一个S13卡上的S3 ASIC。

×绝不（Never）把一个数据单元转发到另外一个S2卡上的S2 ASIC。

因此，这里面所有问题的关键就是：如何将一个新的LCC加入交换平面。

从物理的角度，也可以这样理解，如何将一个新的LCC的S13卡的A0-A2接口接入到一个OIM卡上，而这个OIM卡上，系统中其他LCC的一个 S13卡有相应的光纤接口接入，从而使得系统中所有的LCC可以通过S13卡形成一个连通（交换平面）。否则，系统中就形成一个LCC孤岛。

14. FCC多机交换（5）

通过上节的分析可以得知，如果FCC的一个S2卡单独充当一个交换平面，其相应的OIM卡的J0-J9分布被3个LCC的S13卡的A0-A2接入，CRS-1系统多机互联是无法连接多于（等于）4个LCC机器的。

那么，如何通过一个FCC的24个S2交换卡和相应的24个OIM卡来达到支持多于4个LCC互连?

CRS-1目前推出的一种方案是：多模块交换互联（MultiModule Fabric Connectivity）

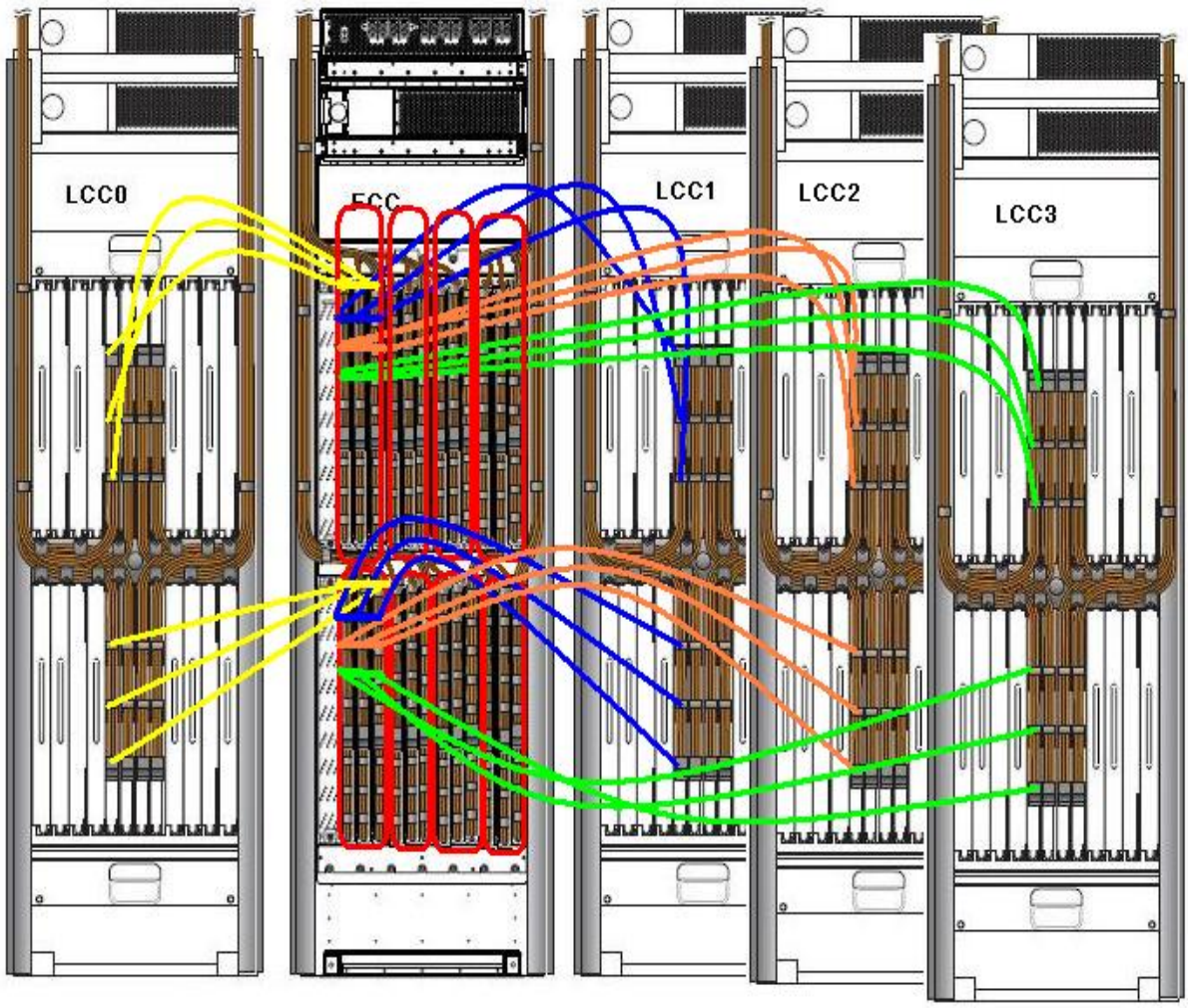
× 通过把24个S2卡和24个OIM卡分组（Zone）的方式。

×24个S2（OIM）分为8个组。每组含有3个S2(OIM）卡。

×每个组相对应一个独立的交换平面。因此，系统支持8个交换平面。

×每个组能连接来自9个LCC的S13卡的A0-A3接口。

上述机制是如何实现的呢？请参阅下图的一个4LCC+1FCC的拓扑。



4LCC + 1FCC互联拓扑

交换平面

从上述图示，我们可以清楚的得出，24个OIM卡被组合为8个区域（Zone）。分别是：
（FCC 上排）OIM0-2，OIM3-5，OIM6-8，OIM9-11
（FCC 下排）OIM12-14，OIM15-17，OIM18-20，OIM21-23
每一组是一个交换平面，因此系统中是8个交换平面。
下面来看LCC的S13卡是如何接入的。
首先我们试图介绍两个概念：S13卡垂直接入 和S13卡水平接入。
S13卡垂直接入：这就是以前讲述的LCC多机互联的接入方式。S13卡的A0，A1和A2分别接入到一个S2/OIM卡的J0-J8中空闲的 接头上。在这种方式下，每个S2卡提供一个交换平面。
S13卡水平接入：这就是本节要阐述的LCC多机互联的接入方式。S13卡的A0，A1和A2分别接入到一个S2/OIM组的Jx上（0<= x <=8）。在这种方式下，3个S2卡提供一个交换平面。
对于S13卡垂直接入，不再在这里解释，通过对前面章节的阅读，应该是非常直观。

对于S13卡水平接入， 从上图所示，可以观察如下特征：
× LCC0的SM0的A0，A1和A2分别被水平的接入到了OIM11，OIM10和OIM9的J0接头上。
× LCC1的SM0的A0，A1和A2分别被水平的接入到了OIM11，OIM10和OIM9的J1接头上。
× LCC2的SM0的A0，A1和A2分别被水平的接入到了OIM11，OIM10和OIM9的J2接头上。
× LCC3的SM0的A0，A1和A2分别被水平的接入到了OIM11，OIM10和OIM9的J3接头上。
通过这样的连接拓扑，4个LCC的S13卡被有机的挂在了一个FCC的交换平面上。从而可以使得，通过这个交换平面，任何一个线卡的数据可以抵达系统中任何一个其他的线卡。
下面来看一下S13卡水平接入的巨大优点：
由于各个S13卡的A0-A2接口是水平接入到一个OIM组中的Jx上。这个x最大是8。
因此，一个OIM组，或者说一个交换平面最大可以同时接入9个LCC的S13卡。8个交换平面，因此能支持8×9=72个S13交换卡。**换言之，这种每个交换平面有3个S2卡组成的拓扑结构，最大能够支持9个LCC互联。**
这也是目前思科CSR-1系统的上限值--8个LCC通过1，2或4个FCC互联。
下面我们探讨思科CRS-1的72个LCC互联的一些理论值问题。
理论问题1：需要多少FCC才能提供其理论上的72个LCC互联？
上述问题的变种是，如何构造一个交换平面，从而，拥有8个交换平面的CRS-1多机系统能把72个16槽LCC互连？
首先，我们来看看CRS-1是否在理论上能支持72个LCC互联，需要多少FCC。
因为，一个LCC拥有8个S13卡。美国S13卡拥有3个光纤接口（A0-A2），
因此，一个LCC需要8×3=24个FCC的OIM接口。
因此，72个全负载LCC需要消耗72×24个OIM接口。
同理，一个全负载FCC拥有24个S2卡和24个OIM卡。每个OIM卡永远9个光纤接口。
因此，一个FCC可以贡献24×9个接口。
因此，为了支持72个LCC互联，所需要的FCC数量= (72个全负载LCC需要消耗72×24个OIM接口) /一个FCC的接口能力
= (72×24)/(24*9)=8。
因此，需要8个FCC，就可以把72个FCC互联起来。
理论问题2：在72个LCC互连的情况下，每个控制平面需要多少个FCC的S2（OIM）交换卡？
72台LCC机器，每个机器拥有8个S13卡。每个机器都必须通过一个S13卡接入到一个 交换平面。因此，一个交换平面需要支持72个S13卡的接入。每个S13卡是3个接口（A0-A2)，因此，是72×3个接口。
因为，每个S2卡的OIM是9个接口。
因此，（72×3） /9 = 24。
因此，在CRS-1的理论最大负载的情况下，每个交换平面需要24个S2卡。
这其实也就是一个完整的FCC的最大S2卡装机量。
换言之，当CRS-1达到其理论最大72LCC互联时，需要8个FCC，并且，每个FCC需要满负载，并且作为一个交换平面 。
读者请注意，从以前的章节中，我们可以做出如下一个总结，

× 如果只为了支持1-3个LCC互联，一个S2卡出任一个交换平面就可以，通过S13卡水平接入方法；
× 如果只为了支持 1-9个LCC互联，需要3个S2卡出任一个交换平面，通过S13卡垂直接入方法。
× 如果只为了支持 72个LCC互联，需要24个S2卡（一个FCC整机）出任一个交换平面。目前不知道如何接入方法，这也是说为什么72机器互联是CRS-1的一个理论值。CRS-1目前只支持前两种规模的互联。
理论问题3： 为什么72是理论最大值？为什么不能通过够买更多的FCC，从而提供80个，88，或100个LCC互联？
我们首先通过反证法来阐述。 假设购买9个FCC，来看看是否能够支持多于73个的LCC互联。
因为每个LCC有而且只有8个S13卡。每个S13卡参加而且只参加一个交换平面，
所以，S13卡的最大离散分布是分别接入到8个FCC的S2（OIM）卡上。
现在我们先将前72个LCC通过上述方案接入。 我们得到如下：
第1-8个FCC的24个S2（OIM)卡的J0-J8接口都全部占满。72个LCC的S13卡也全部用光。
现在还剩下一个LCC和一个FCC。
如何连接？这第73个LCC无法接入到任何一个交换平面中，无法加入路由器系统。

15. 参考文献

思科CRS-1主页：CRS-1 Homepage

思科CRS-1产品系列：Models Comparison

思科CRS-1 16-Slot Single-Shelf System

思科CRS-1 8-Slot Single-Shelf System

思科CRS-1 Modular Services Card (LC)

思科CRS-1 4-Slot Single-Shelf System

思科CRS-1 24-Slot Fabric-Card Chassis