



面向未来的思科 8000 系列 路由器平台

赵勇 / Frank Zhao

思科顾问工程师

May 2020

议题

- Cisco 8000
- Cisco Silicon One™
- 光模块
- IOS XR7
- 应用案例
- Sonic on Cisco 8000

Cisco 8000

为什么需要新的 ASIC 架构？

- 需求在改变
 - 更加关注带宽和功耗
 - 超大规模云
- 底层组件技术的变化
 - SerDes, SRAMs, DRAMs, Silicon processes, Optics
- ASIC架构通常持续约10-15年



8000 Series 位置

- 带宽需求 5-200+ Tbps
- 首先会在核心和汇聚
- 选择 web-scale data center roles
- 100-400 Gigabit Ethernet
- IOS XR

Cisco 8200 和 8800 路由器

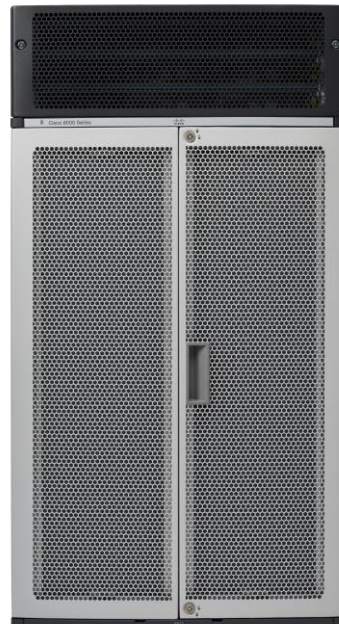
8201



8202



8808

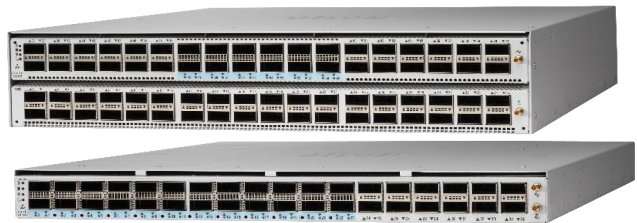
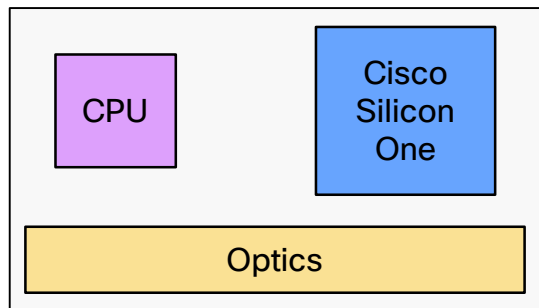


8812

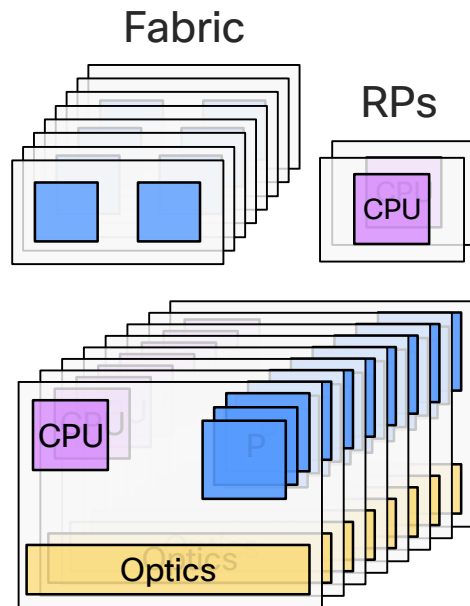


8818

一个 ASIC - 两种架构



8200 Series

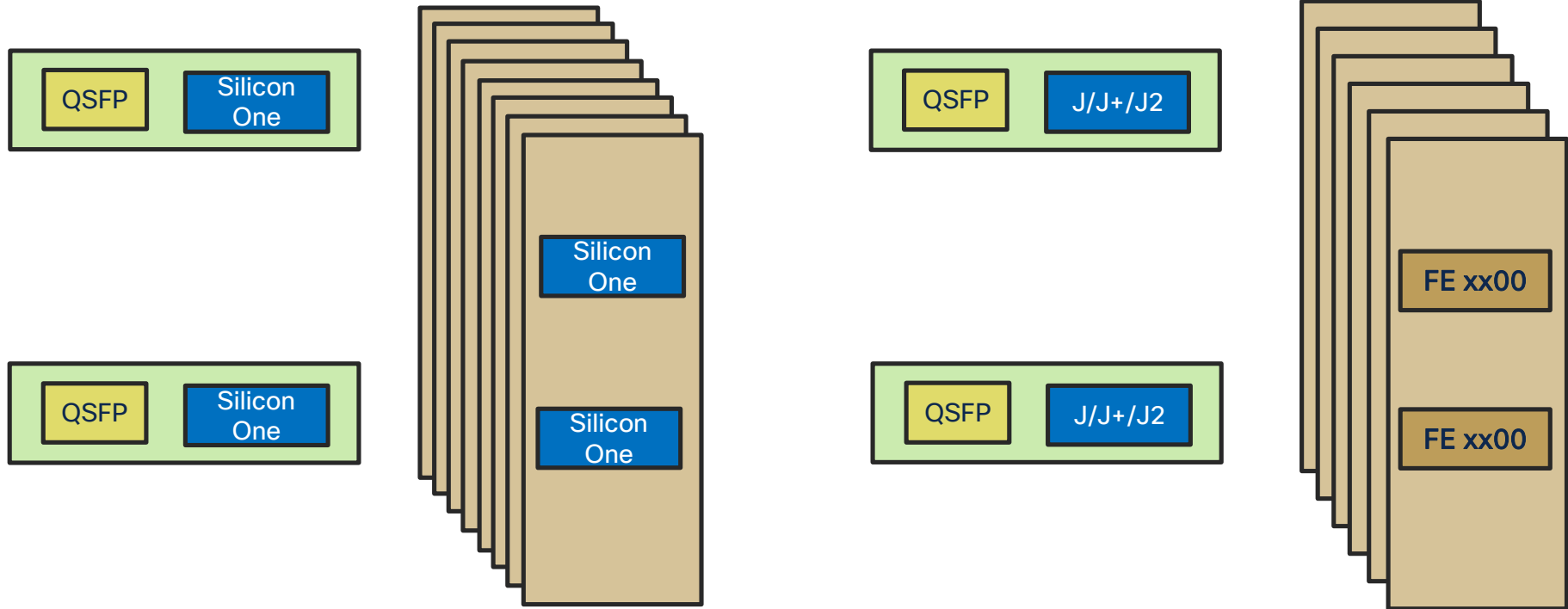


Line Cards



8800 Series

统一的 NPU

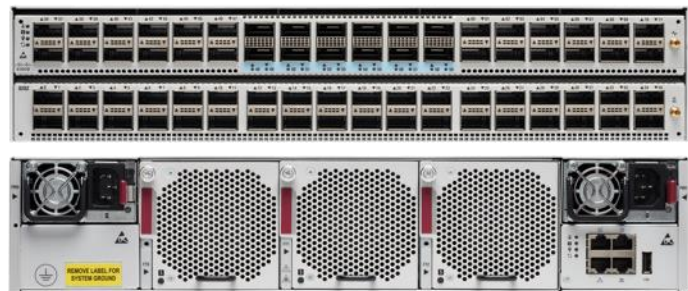


Cisco 8200 系列

固定型号



Cisco 8201
24x QSFP56-DD + 12x QSFP28



Cisco 8202
12x QSFP56-DD + 60x QSFP28

- System on a Chip (SoC) 模式
 - 10.8 Tbps 分配给到网络端口
- 单一ASIC设计实现简单高效的散热
- 结合使用QSFP56-DD和QSFP28以充分利用带宽
- 固定型号每RU配置36个端口
- 双向通风
- 8201: 4w/100g; 8202: 7w/100g

Cisco 8800 模块化系统

8808, 8812, & 8818通用组件

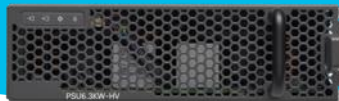


高达 8块矩阵卡



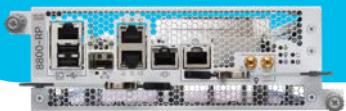
正交设计

电源



- 60A 48V DC (4.2kW)
- 100A 48V DC (6.3 kW)
- AC & High Voltage DC (6.3 kW)

路由引擎(x2)



- 4-core Broadwell
- 32 GB (Default) to 128 GB DRAM

线卡

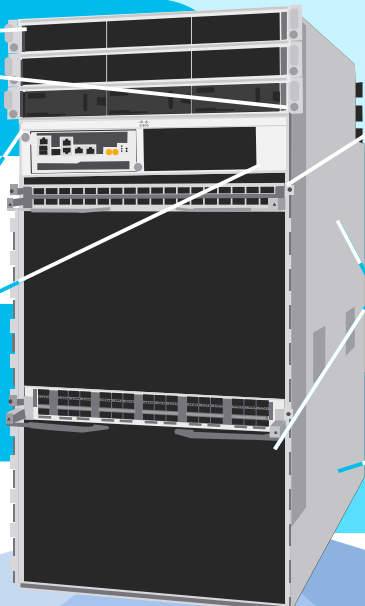


36x 400G QSFP56-DD



48x 100G QSFP28

实现EOBC小中背板
(10G 从RP到LC)



每100GbE耗电11瓦



8808



8812



8818

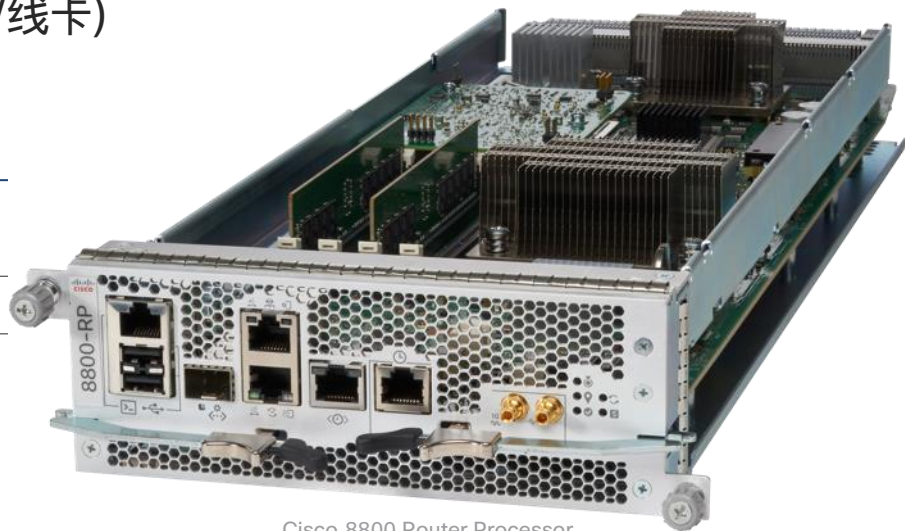
Cisco 8800 系列路由引擎

8800 RP

- 所有Cisco 8000模块化系统之间通用
- 无需连接到矩阵(EOBC小中板，10G带宽/线卡)

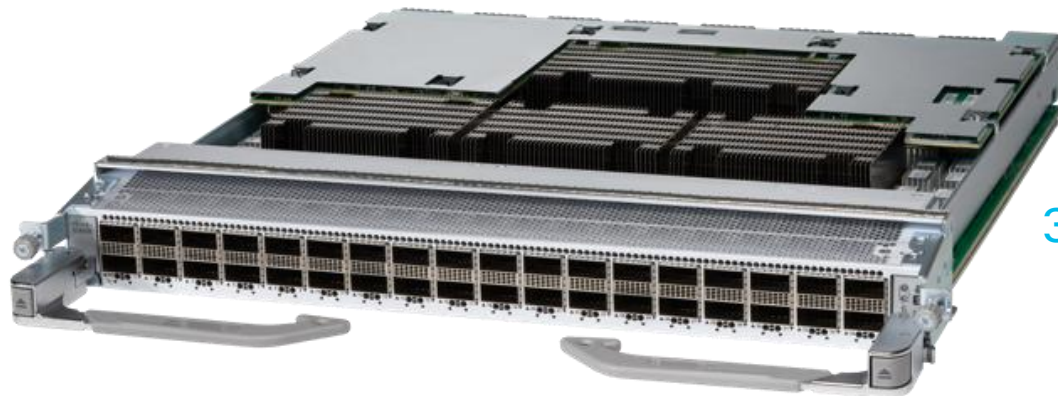
RP 8800 特性

端口	<ul style="list-style-type: none">• BMC & 管理以太网口• 灵活的2x USB 2.0 1A 端口
处理器	<ul style="list-style-type: none">• 4核 2.4 GHz Broadwell CPU
其它特性	<ul style="list-style-type: none">• Timing Class B• IEEE 1588• Sync 0• TOD• 10 MHZ / 1 PPS



Cisco 8800 Router Processor

Cisco 8800 系列 线卡



36x400G QSFP56-DD



48x100G QSFP28

Cisco 8800 模块化系统

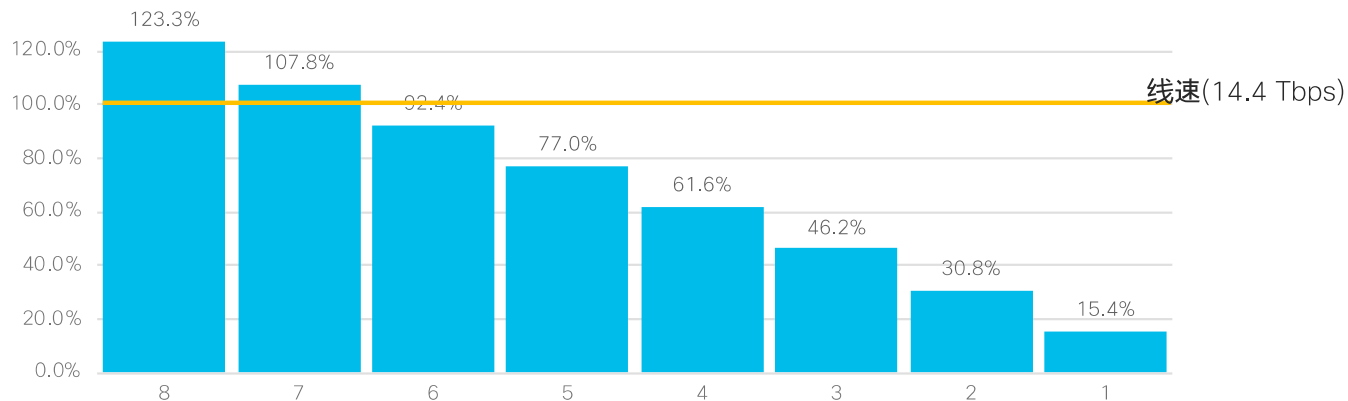
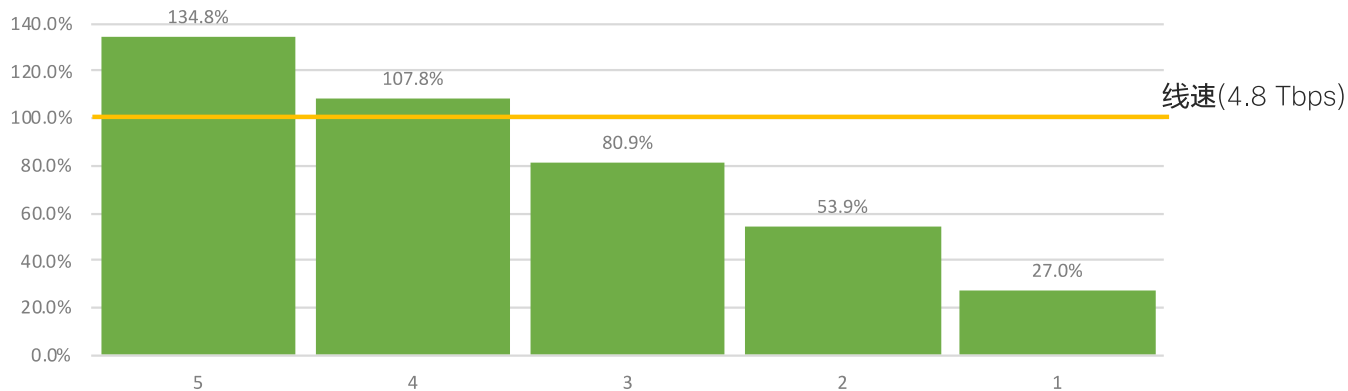
交换矩阵

- 正交直连消除了对中板和/或背板的需求。
- 线卡和风扇组之间最多有8个矩阵卡
 - 48x100G线卡需要5个矩阵卡实现N+1冗余
 - 36x400G线卡则需要8个矩阵卡，以实现N+1冗余
- 矩阵模式下的Cisco Silicon ONE Q100 ASIC
- 矩阵卡特定于每个机箱，垂直安装在线卡后面



Cisco 8000 系列

5/8FC模式下，矩阵的灵活性



Cisco 8800 系列的冷却系统

Cisco 8808, 8812, 8818 风扇

位于卡上方和光模块
之间的进气口

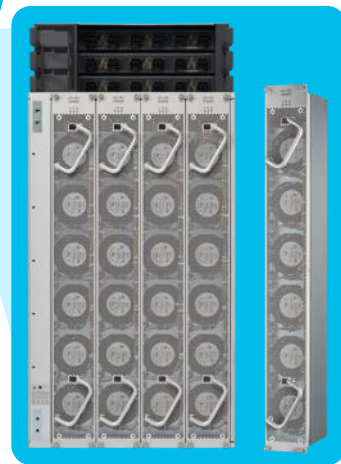


不带门和过滤器的Cisco 8812



带有门和过滤器的Cisco 8812

过滤器可选



4 风扇组 - 特定于每一个机箱

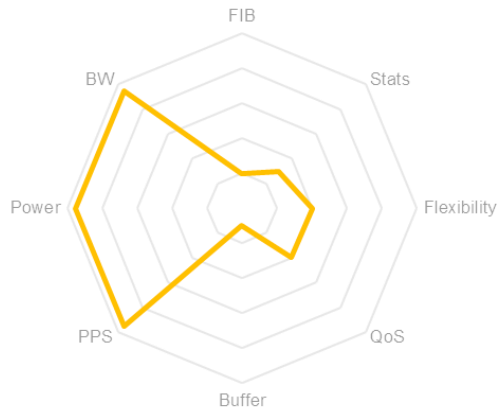
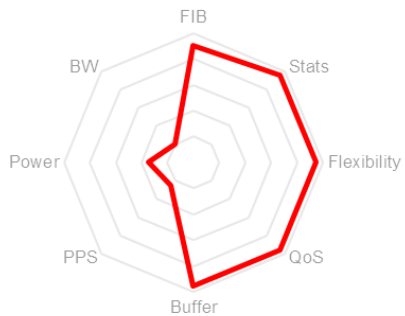
Cisco Silicon One™

Cisco Silicon One

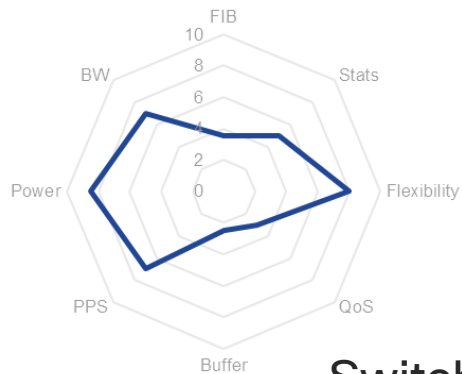
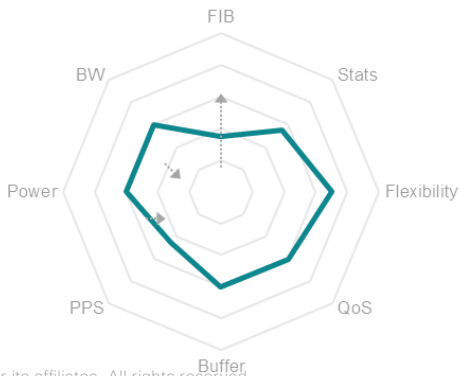
- 融合了路由和交换的能力
- Clean sheet design
- 广泛的带宽和性能优化
- 第一个设备是 10.8 Tbps Q100
- 单芯片担任多个角色
 - Router on Chip
 - Line card NPU
 - Switch fabric



路由 vs. 交换 ASICs

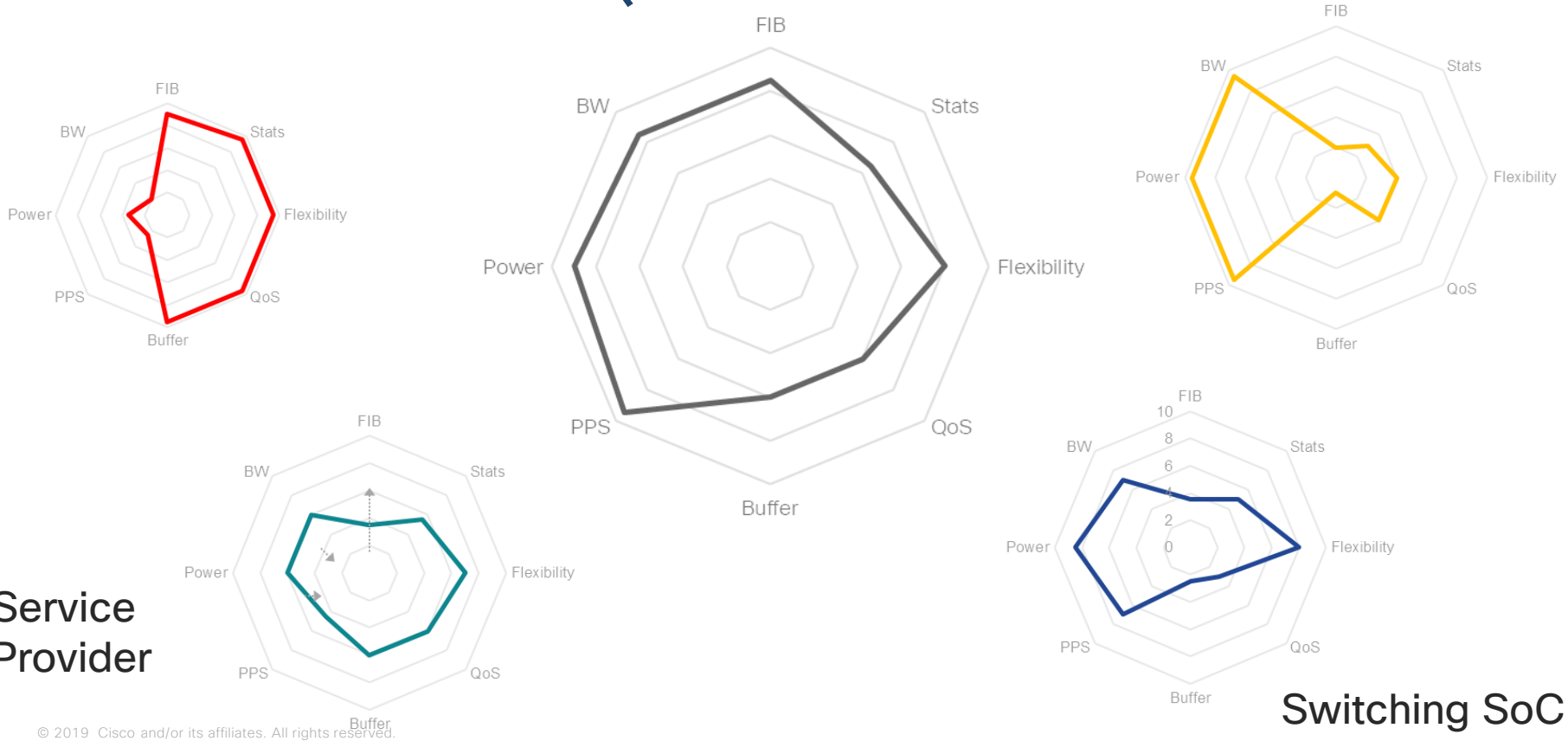


Service
Provider



Switching SoC

Cisco Silicon One Q100

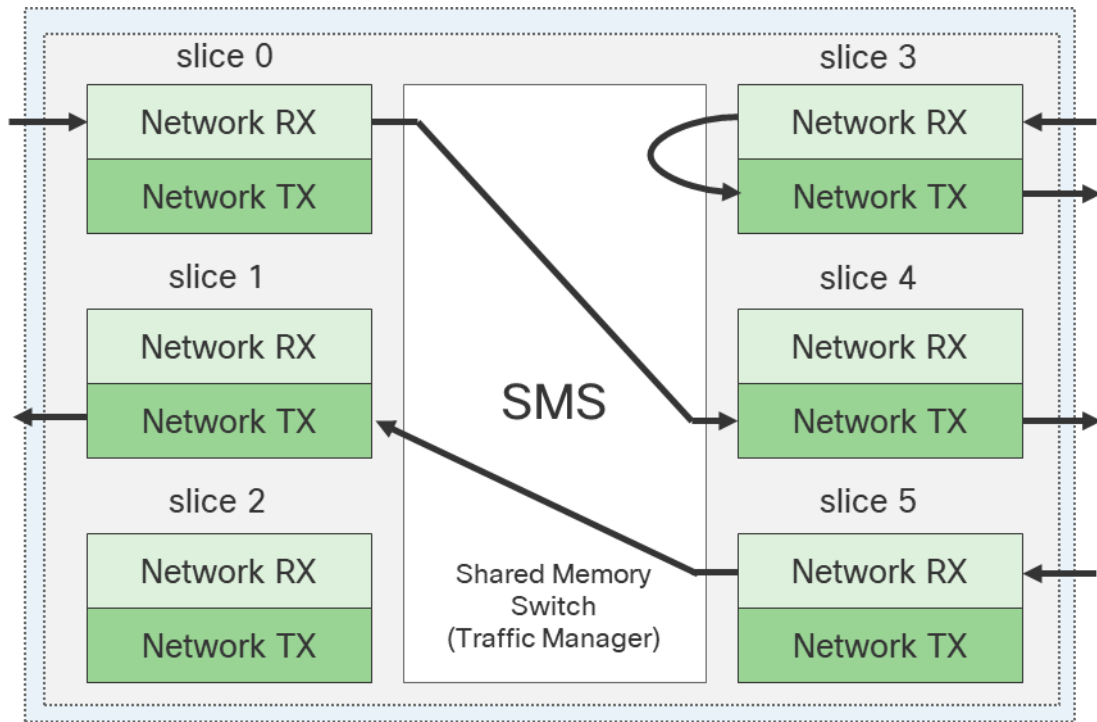


芯片创新

- 业界最佳的硅芯片团队
- 消除片外存储器
- 片上 HBM (High Bandwidth Memory) ， 用于 FIB 和大缓存
- 支持6B pps以上的灵活RTC引擎
- 先进的数据结构和查找算法
- 可扩展， 灵活的 multi-slice 架构



芯片的“盒式设备”模式



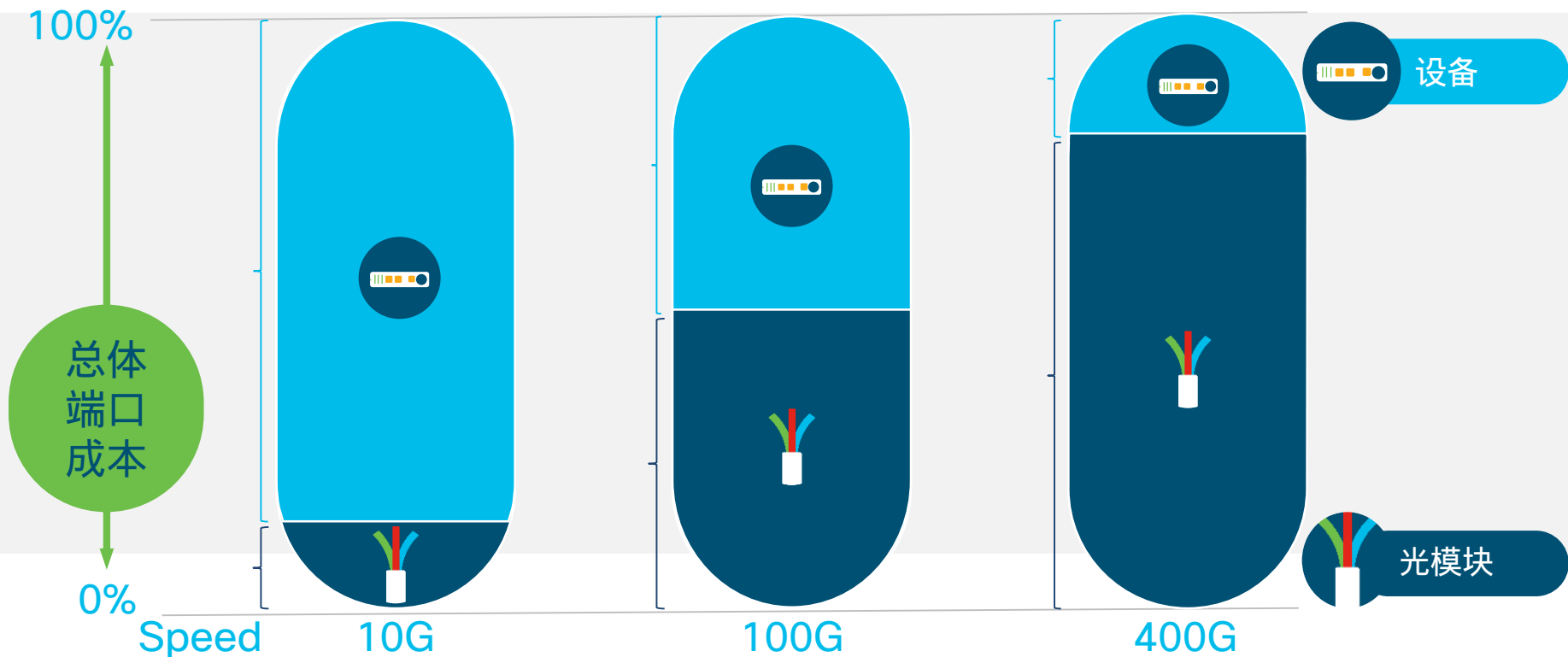
硬件的挑战

- 大多数路由器组件不遵循摩尔定律
- 从2016年到2019年尝试将系统带宽提高4倍
- 广泛的平台挑战 - 不只是新的 ASIC

- 内存
- 内部带宽
- 电力分发
- 热管理
- 高密度 optics

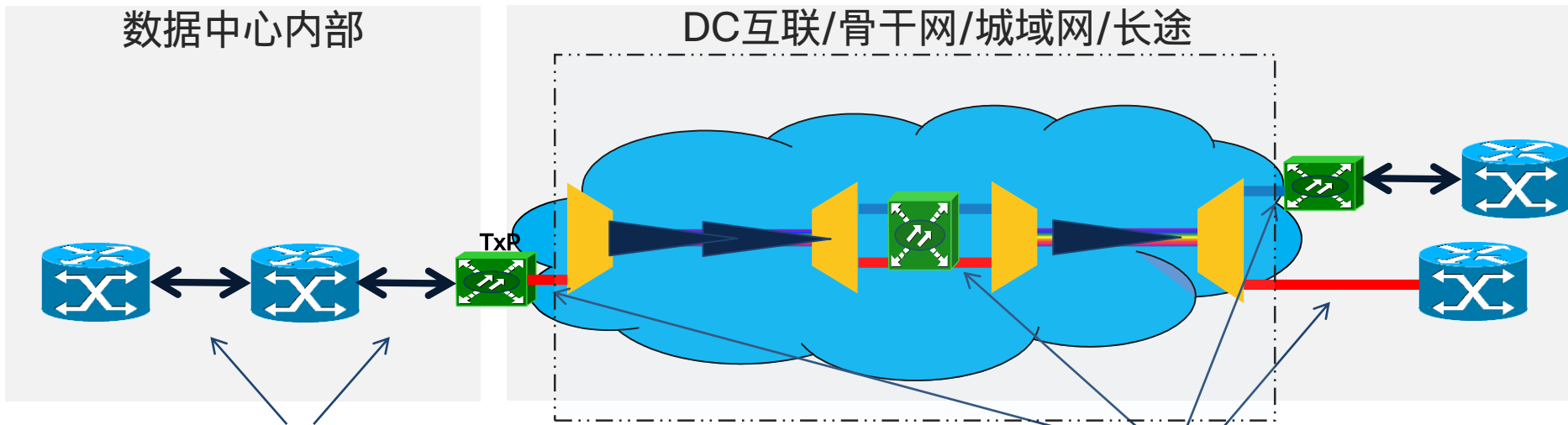
Optics 光模块

随着端口密度/速率的提高，光器件成本占比越来越高



为什么？对比光器件，端口成本降低的速度更快
光器件复杂性随速率的增加而增加

思科在光器件领域的关注点和持续投资



DC & Client Optics (Short Reach)

LIGHTWIRE

Acquired, 2012
\$271M

LUXTERA

Acquired, 2019
\$660M



Optical Systems

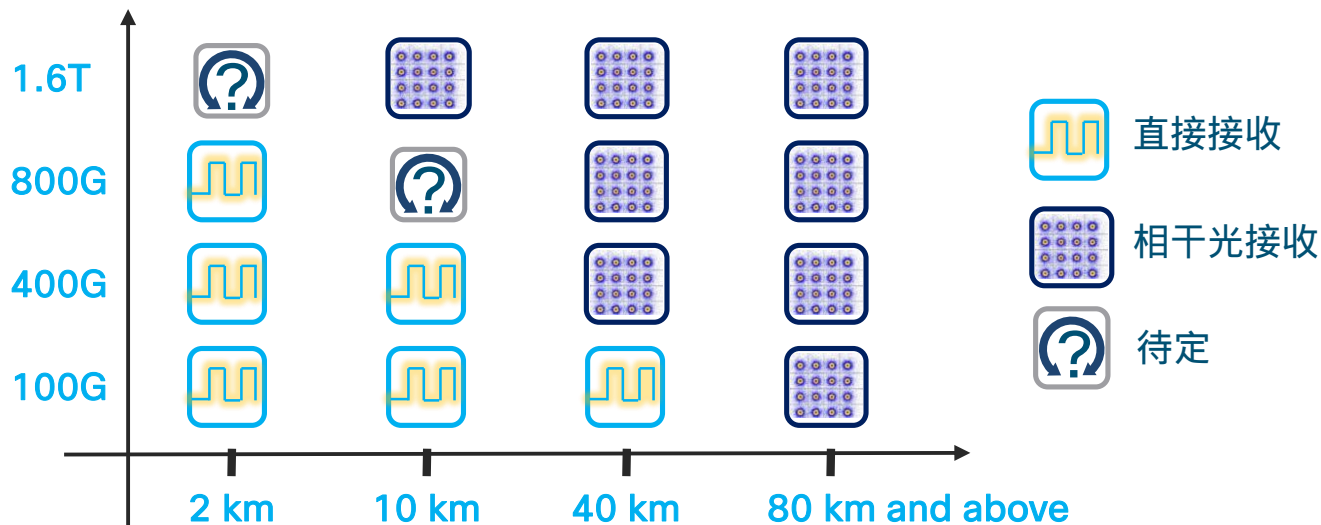
Core Optics

Acquired, 2010
\$99M

ACACIA COMMUNICATIONS, INC.

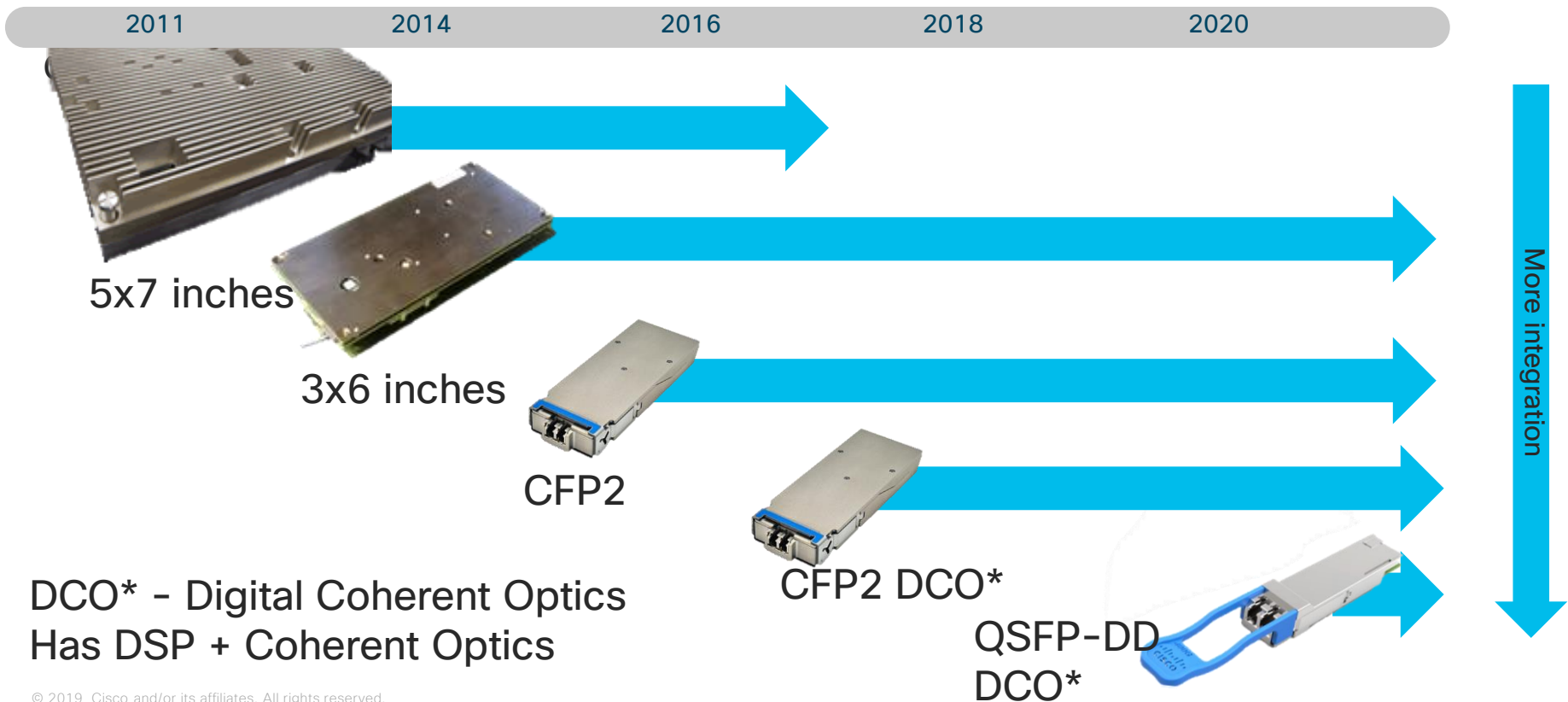
Announced, 2019
\$2.6B

相干检测技术成为高速光通信的核心技术



随着光模块速率提高，“相干检测”技术将逐步替代“直接检测”技术

相干检测技术的演进



相干100/200GE 80公里光纤直连解决方案

- CFP2-ACO – only for 6xDWDM LC

- Coherent DWDM - Analog
- Requires host DSP for IPoDWDM
- 100G QPSK target 4200km
- 150G 8QAM target 1800km
- 200G 16QAM target 1000km
- Dark fiber target 80km



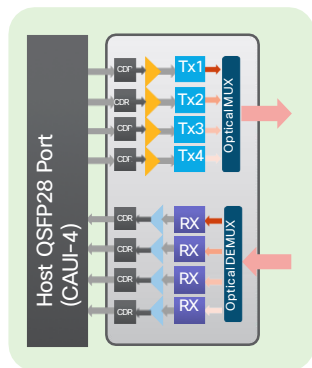
- CFP2-DCO – only for 1x/2xCFP2 MPA

- Coherent DWDM - Digital
- Integrated DSP for IPoDWDM
- TOF and non-TOF version
- 100G QPSK target 2300km
- 200G 8QAM target 700km
- 200G 16QAM target 400km
- Dark fiber target 80km
- Staircase FEC version 100G only with TOF

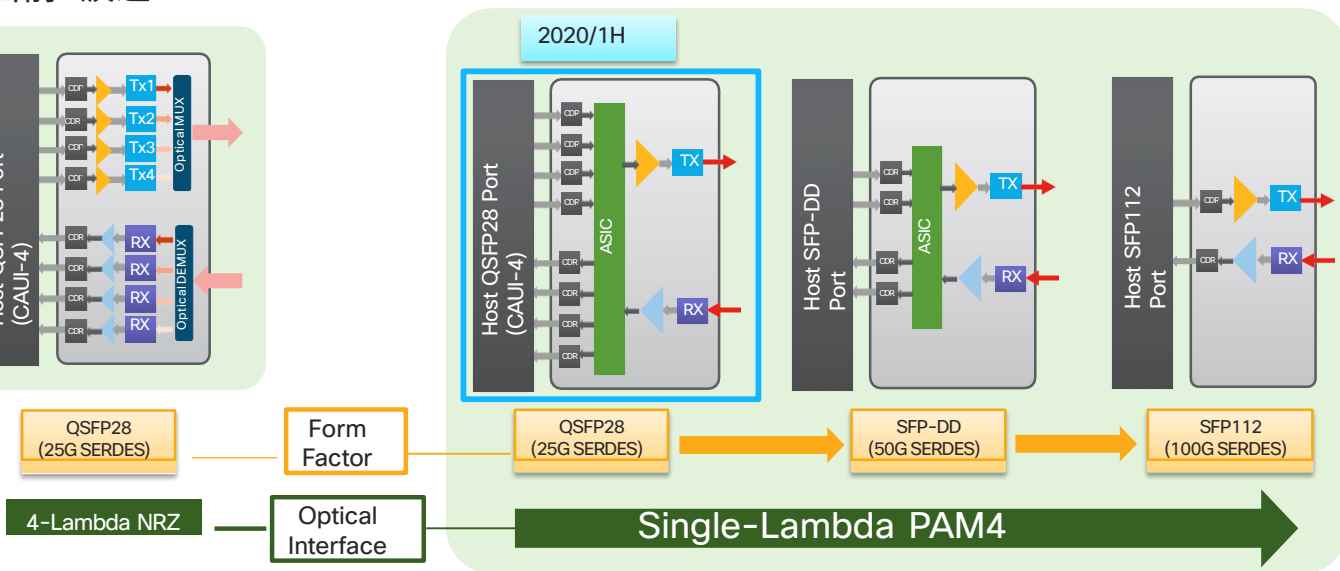


单波100GE低成本光模块的演进

当前4波道100GE



单波长100GE (预计2020年上半年商用, 价格大幅度下降)



软件 IOS XR 7

Cisco IOS XR 7

重新定义软件以实现更好的操作



简单

- 经过优化以减少内存，下载和启动时间
- 使用 SR/EVPN 简化协议
- 安全的零接触部署



时尚

- 开放的 APIs
- 可定制的软件映像
- 云增强



值得信赖

- 在启动和运行时评估硬件和软件的真实性和完整性
- 实时了解信任状态



50% 更少的内存占用



50% 更快的开机速度



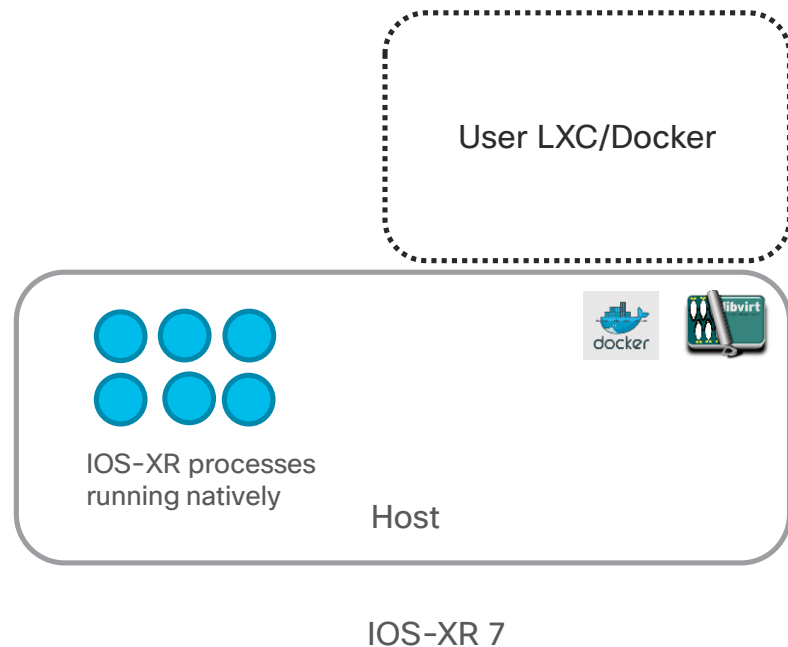
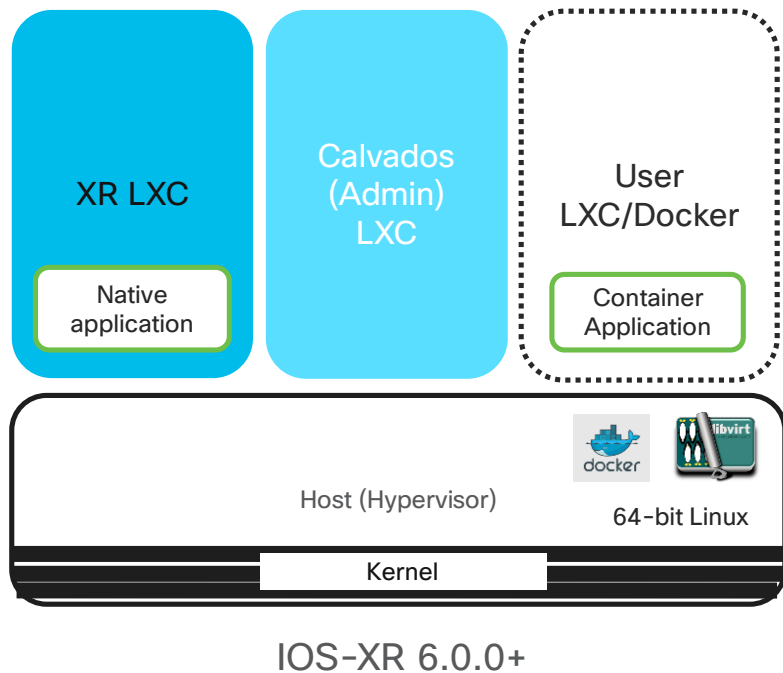
40% 更小的镜像大小



40% 更快的下载

Cisco IOS XR 7

重新定义软件以实现更好的操作



Ops Benefits

- File movement between containers no longer needed
- Access (AAA) to different parts simpler – ZTP, Scripting

IOS XR 设备可编程性的基础

IOS XR



简单

传输简化

使用Segment Routing简化传输

服务简化

EVPN 实现统一的服务交付, 云原生BNG

运营简化

NSO, Yang Development Kit, Yang Suite



时尚

先进的telemetry

启用分析和机器学习的可见性

软件中各层嵌入的API

通过模型驱动的可编程性实现创造力和灵活性

可以与第三方软件兼容

使用目标第三方软件补充IOS XR基础



值得信赖

信赖始于硬件

防伪和信任锚基础结构

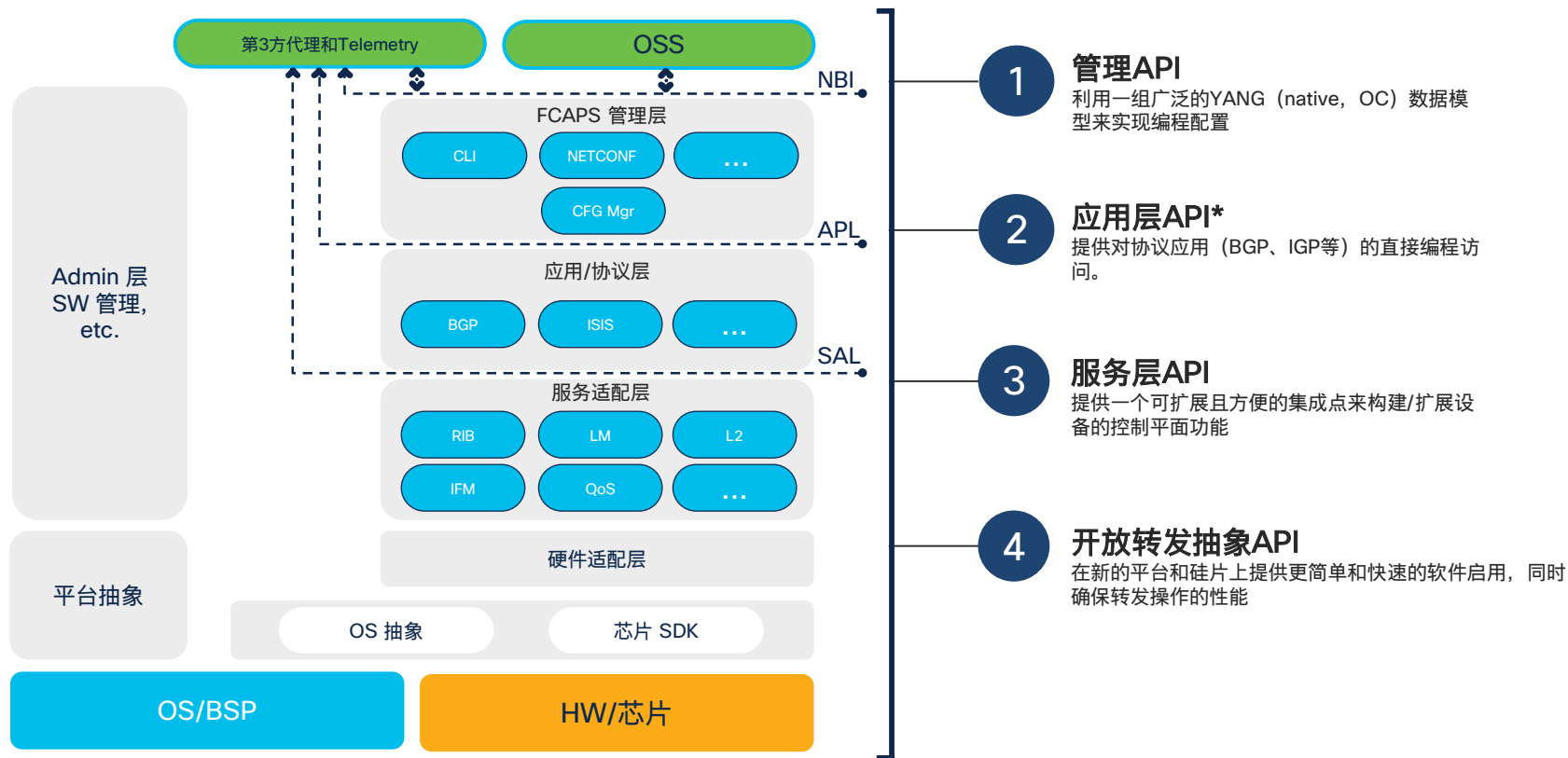
一个值得信任的网络操作系统

图像签名和安全启动基础设施

运行时值得信赖

运行时防御, 加密传输, DDoS保护

API-驱动, 多层软件架构

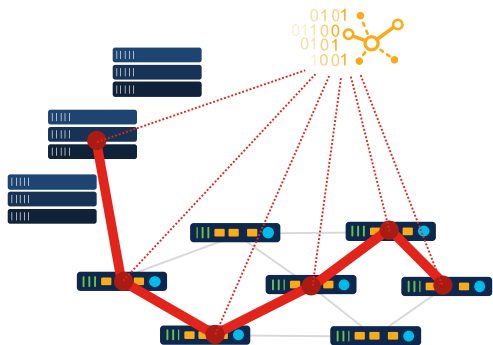


* The use case & implementation of the Application Layer API is currently under evaluation & external availability of the API is TBD

服务层API

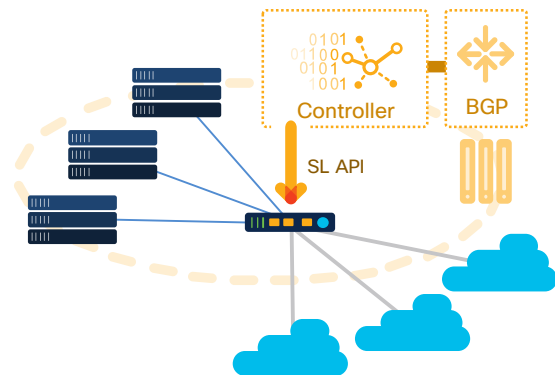
使用案例

流量工程和路径选择



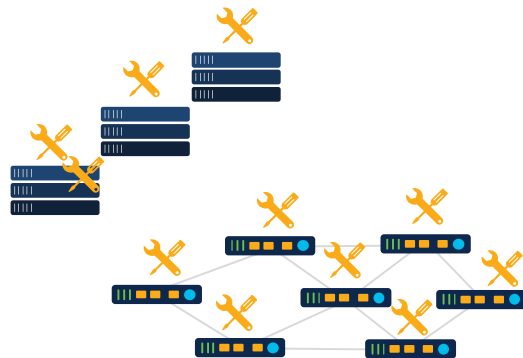
通过路由/标签操作为应用程序设计路径，所有这些都基于用户特定的逻辑

可编程路由下载



可编程路由下载到CDN PoP路由器，以优化TCAM空间

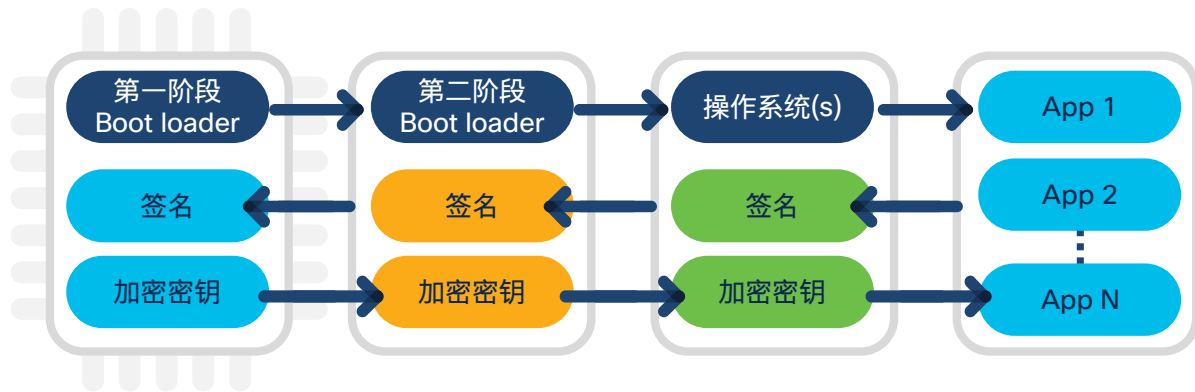
自带的协议/代理



现成的代理以及自定义协议，可与标准协议共存以影响路由

建立信任

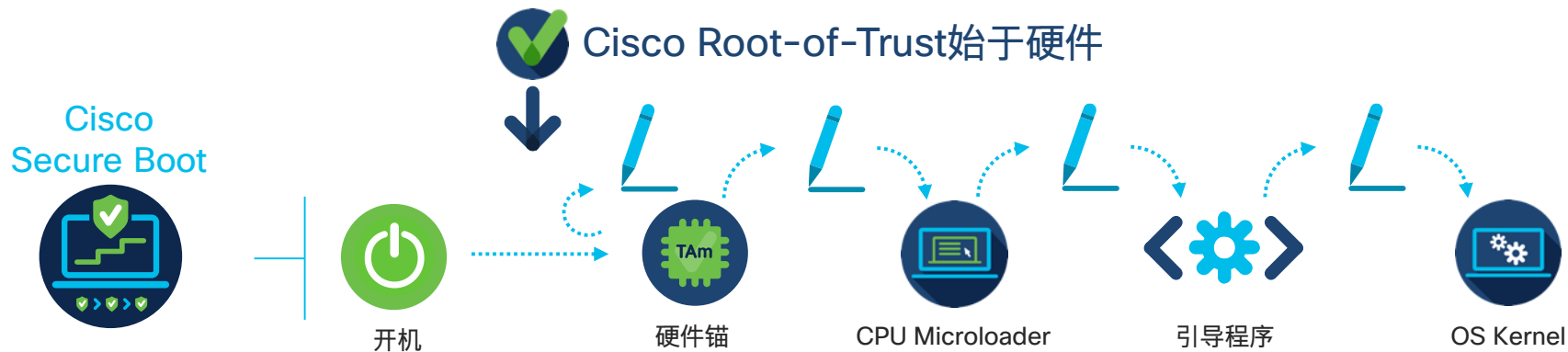
“我怎么知道我可以信任路由器的硬件、固件和软件？”



- 信任链
- 在运行代码前检查签名(Signature)
- 测量从硬件直到OS

建立信任

安全启动



云增强的可靠平台

重定义网络运营

最新: Cisco Crosswork Trust Insights / Data Gateway



4

完整性可见度 (Boot & Run-time)

3

运行时仍然保持信任

最新: 运行程序/进程的指纹和完整性强制应用

2

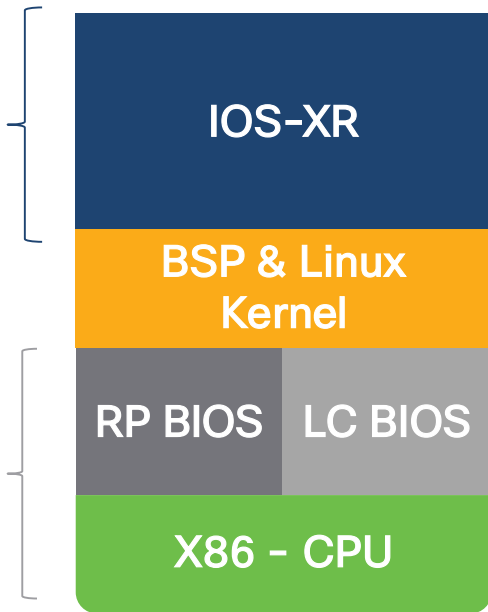
建立可被信任的 NOS

最新: 基于思科芯片级保护& Image& Package 签名的安全启动

1

建立基于硬件的信任

最新: 思科芯片级保护和存储



- 外部信任态势评估
- 完整性度量(Integrity Measurement) 的安全引用

- Run-time 防御
- 完整性度量架构
- 启动完整性可见度

- 地址空间布局随机性
- 安全增强的Linux

- 硬件完整性度量
- 安全启动

- 信任锚模块(TAM)
- 安全 JTAG
- 安全UDI

应用案例

User1 SP Aggregation Roles

- Position A: 8808/8812, 8202
 - Spine and Leaf cluster
 - Spines 连接到Leaf, PE, Core router
 - Leafs 连接到接入汇聚 B
- Position B: 8802
 - 两个汇聚box, 连接A
- Features:
 - L3VPN option C
 - BGP PIC Edge w/ Multipath
 - BGP FS
 - L2VPN

User2 Core and LSR

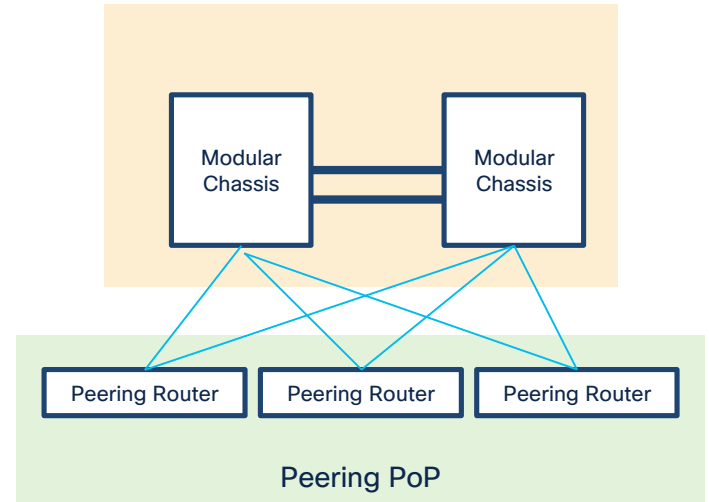
- Core “P” Router with 14.4T /slot: 8812
- 100G/400G Spine LSR router: 8201
- High density 100G/400G
- Significant lower power consumption
- Big user of SR and will deploy SRv6 next year

User3 Core Router

- Core Router with 14.4T /slot: 8818
- ZR/ZR+ optics support on 400G
- Cisco One ASIC (scales to 25TB per slot)
- Significant lower power consumption
- Timing support
- SR & LDP, SR-TE & RSVP-TE inter-working

User4 Peering Position

- Peering role
- 1RU fixed chassis, high density: 8201



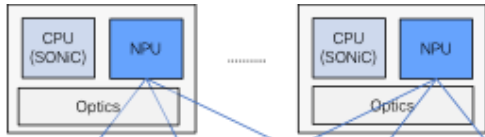
SONiC on Cisco 8000

SONiC on Cisco 8201

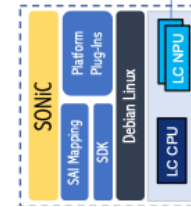
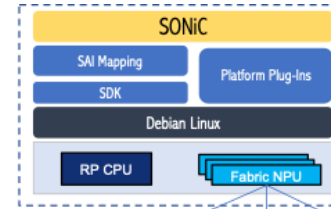
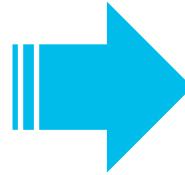
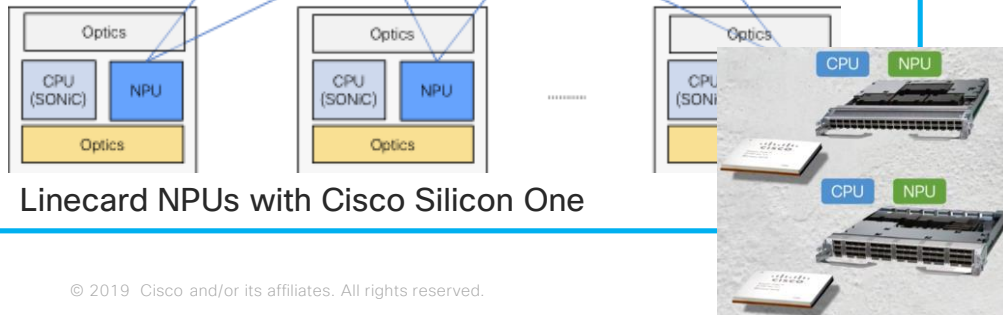


SONiC on Cisco 88xx Modular Chassis

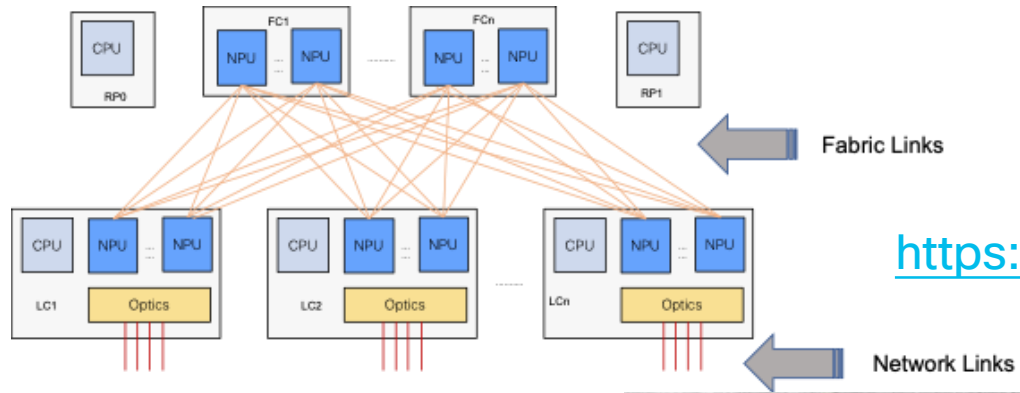
Fabric NPUs with Cisco Silicon One



Linecard NPUs with Cisco Silicon One



SONiC on Cisco 88xx Modular Chassis



<https://xrdocs.io/ocp/>

