


# 基于SDN-OpenFlow的网络智能控制

演讲人：Min Luo (罗敏) / [min.ch.luo@huawei.com](mailto:min.ch.luo@huawei.com)

职 位：未来网络研发主管兼首席架构师

香农实验室, 华为技术有限公司



重构网络 敏捷未来

Best Time to Change Redefine the Future

# 目录



**1** SDN集中中央管理使智能控制优化成为现实

**2** 智能控制必须广义-支持异构网络/设备/协议

**3** 智能控制系统应该可编程，自适用，自管理

**4** 智能控制信息来自边缘-网络与应用集成点

**5** 华为香农实验室SDN-OF技术研发

# 中央集中管控降低管理复杂性以及日益增加的技术创新的难度

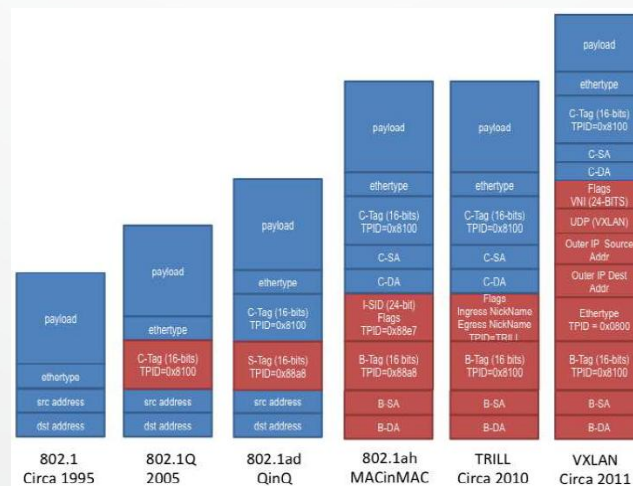


- 挑战:

- 提供先进的网络基础架构和服务，同时降低TCO
- 将现有物理网络基础架构的智能和控制由分布在网元中的硬件转移到运行在通用服务器中的软件，以使得各种网络服务更加智能，快速和便宜
- 灵活，可扩展，高可靠，自适用，模块化的控制和转发平台

- 现实：

- 异构的网络设备和架构
- 数千不同层次，不同作用的协议：大部分不能互通，但很多交互，没有开放式的接口标准
- 分布式的控制模式无法达成全网资源的充分利用，以及对应用QoS的保障
- 动态的网络状态，而配置通常是固定，人工完成的，导致网络集成管理程度差
- 网络配置复杂，低端，且易产生各种错误并衍生安全问题: 成百万行的配置脚本，成千上万的不同设备，每天成百上千的各种需求或环境变化。
- 缺乏创新的能力来适应日益增长的新的业务需求，控制成本
- 扩容往往解决不了堵塞问题，反而导致问题更趋于复杂，资源利用率更低。



# 抽象+层次：解决复杂问题的通用途径



- 统一的方法或模型来思考网络中的“网元黑盒”
  - 路由器，交换机，防火墙，NATs，负载均衡器，...
- 集中统一的网络视图
  - 简化控制逻辑和相关操作
- 剥离策略与配置机制
  - SDN-OF并不直接实现，但使得这种分离成为可能
- 控制与转发平面的抽象
  - 转发：Openflow以及流（flow）空间的抽象
  - 控制：全网视图（物理，逻辑）
    - 分布式全网状态的“逻辑集中”
    - 分布式控制机制的“逻辑集中”

# SDN解决之道 — 抽象层次化的控制 转发分离



- **SDN基本思想：**控制面与转发面分离，控制面对转发面进行集中控制，转发平面设备依据控制平面下发的路由策略实现数据转发
  - 集中的控制平面根据全网的视图，动态、灵活、高效地管理和配置网络
  - 增强运营商对网络的可管理性和控制力度
  - 方便对网络基础设施进行定制和优化
- **集中的SDN网络控制器**
  - 控制器北向接口可以对接多种管理应用（多路径算法，资源调度算法等）
  - 基于控制器维护的全网拓扑、状态、流量特征和统计信息，北向管理应用可以实现路径选择和资源调度等的优化
- **控制和数据转发的分离有效地屏蔽了底层不同硬件交换设备的差异，使得统一管控由不同设备组成的异构网络成为可能**

# SDN-OpenFlow集中管控的价值



- 端到端的满足各种业务QoS需求, 基于全网拓扑状态以及流量特征的全局最佳路由
- 动态灵活地管理调度网络资源, 从而大大提高资源的利用率 (Google: 30% → 90%+)
- 更加容易地以软件创新的速度来开发和部署新的增值业务
- 更加灵活和可靠的路由技术: 负载均衡及自动备份容错等
  - 多路径 vs 单一路径:
  - 具有QoS特征和保证的路由
  - 能有效应用于大型网络 (运营商, 跨域DCN。。。)
  - 用户可选可配的路由
  - 更好的基于服(业)务感知的资源控制: 流量到资源的服务映射; 动态配置新的服务逻辑



# 集中控制下的新能力



- 智能网络控制和管理
  - 自动网络拓扑发现
  - 交换机功能发现，协议版本协商，学习和发现
  - 智能混合组网
  - 智能链路或其他故障的自我修复
  - 负载均衡
- 基于全局拓扑、状态及流量特征的网络优化
  - 非-最短路径（SPF）的（多）路由计算
  - SDN-TE
  - 网络扁平化
- 全网资源智能管理

# 目录



- 1 SDN集中中央管理使智能控制优化成为现实
- 2 智能控制必须广义-支持异构网络/设备/协议
- 3 智能控制系统应该可编程，自适用，自管理
- 4 智能控制信息来自边缘-网络与应用集成点
- 5 华为香农实验室SDN-OF技术研发



# SDN-OpenFlow 协议的演进使设计实现 日趋困难



- OF1.0, 2009年12月: 实验性, 简单, 但缺乏支持企业所需的功能以及大规模部署对性能, 可靠性等的要求
- OF1.1, 2011年2月: 多流表, 组表, MPLS和VLAN支持
- OF1.2, 2011年12月: 多控制器, IPv6 支持
- OF1.3, 2012年4月: 辅助通道, 流量计数及控制, 带宽保证等; 基于TLV嵌套式多部分查询
- OF1.4: 同步表, 流监控, 组合消息, 。 。 。 ?
- OF2.0: 更加抽象的控制转发分离模型, 。 。 。 ?

# 不兼容的南向API可能导致网络分裂以及用户接受采用的难度



- 各种新的功能带来了新的代价，尤其“重生”的网络复杂性以及有效实现这些功能在架构和组件设计实现上的难度，OF协议本身从1.0到1.1，1.2都产生“革命性”变化：
  - OF1.1在2011年6月公布，但直到2012年10月，支持1.1/1.2的控制器和硬件交换机才第一次宣布（华为香农），并通过ONF的PlugFest测试。
  - 2013年2月欧盟EANTC测试再次验证后，EANTC在3月发布的白皮书以及SDN Summit/Ethernet World大会上公开宣布了上述事实
  - 2013年6月ONF PlugFest终于有6家厂商宣布支持OF1.3的交换机，但除了2家设备能通过基本的L2/L3测试以外，只有一家（华为香农）能实现OF1.3的核心功能（如带宽保证，流量控制等）
- 过去相对比较稳定的数据转发面也因快速变化的南向OF协议变得更加动态，一些真正能实现SDN基本价值的功能（如多流表）迟迟不能应用
- 不兼容的南向API可能造成网络分裂问题，也增加了用户选择使用SDN-OF技术的困惑
  - 重复投资，产品寿命短且不可持续
  - 从支持一个协议到新的协议的网络技术不是简单的迁移，必需全部重建！

# 广义智能SDN-OF网络控制器



- 真正能帮助推进SDN的发展和接受的是构建一个广义SDN网络控制器
  - 能够控制管理不同（甚至版本不兼容）的交换机混合组网(OF1.0和OF1.2不兼容)
  - 能够快速适用新型网络架构和协议的变更
  - 能够融合异构（IP, OF, Optical)网络的互通，配置和管理

# 目录



- 1 SDN集中中央管理使智能控制优化成为现实
- 2 智能控制必须广义-支持异构网络/设备/协议
- 3 智能控制系统应该可编程，自适用，自管理
- 4 智能控制信息来自边缘-网络与应用集成点
- 5 华为香农实验室SDN-OF技术研发

# pCONF -下一代可编程，自配置，网络 (元) 模型驱动和应用框架平台



- 为了适用快速变化的网络协议（如OpenFlow），中央控制器必须采用一个可扩展，模式化，可管理，高可靠的SDN软件架构
- 必须有效地采用和融合自80年代中期软件危机和工程化以来的架构和设计原则，最佳实践，以及近年来在分布式计算，智能分析与控制，和可重用软件框架方面的最新进展
  - 模式驱动（MDA）软件架构技术和基于MDA的网络控制器架构设计技术
  - 采用可编程技术实现动态集中可控可配置的策略来支持“软件”和开放式的网络节点和管理应用
  - 各种模型能够用一种统一的元模型和支持语言来描述，并且采用可配置参数化的技术来实现未来网络的体系结构和满足不断发展的协议和扩展要求
  - 模型可以实例化
  - 尽可能重用已有技术

# 基于虚拟化的层次结构



- 虚拟化将在这种架构中发挥更重要的作用
  - 节点的虚拟化（包括存储和计算资源）
  - 网络（带宽，地址等）资源也需要更进一步的虚拟化
  - 统一管理所有这些资源的集成管理应用和共享的智能信息
- 支持并行处理模式和实例化
- 实例化的pCONF可以是一个可以控制现行的基于IPv4的Internet或基于OF1.x的一个混合网络，亦或多个异构网融合的大型网络



# 智能零配置自动管理



- 目标：敏捷，灵活，可管理
- 自我管理：自我配置，优化，修复和防护
- 支持各种自我管理的特性和自适用的运行模式，这样网络总是（尽可能）自动平稳并运行在近乎优化的状态
  - 通过动态可选可配的网络管理平面实现对网络（物理或虚拟）的有效动态管控
  - 网络管理和编排的自动化：配置过程可由7天将为30秒！
  - 通过获取知识来建立分布式智能库和管理平面协作式的智能决策
  - 通过策略来设置高级管理目标，实现内置管理职能的有效协作

# SDN网络物理分布逻辑集中多控制器



- 实现基于部分网络状态/视野的分布式决策
  - 稳定性和快速收敛
- 融合网络内大多数管理平面功能，真正实现“集中管控”
- 网络控制器成为可能的单点瓶颈：
  - 可动态分布配置的，强大的中央控制器集群：  
物理分布，逻辑集中
    - 性能 → 扩展性
    - 可靠性
    - 安全性

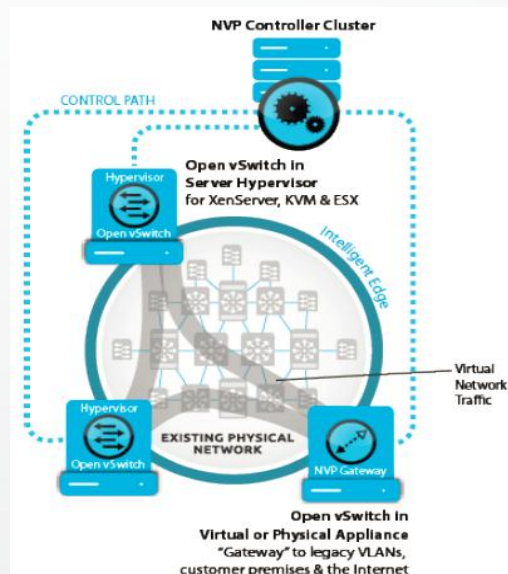
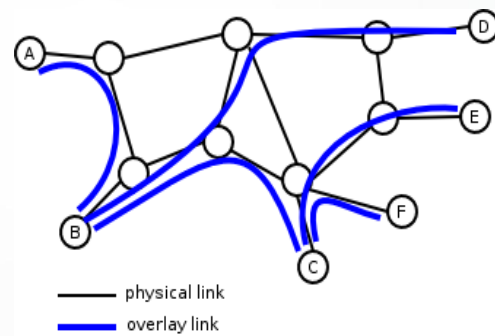
# 目录



- 1 SDN集中中央管理使智能控制优化成为现实
- 2 智能控制必须广义-支持异构网络/设备/协议
- 3 智能控制系统应该可编程，自适用，自管理
- 4 智能控制信息来自边缘-网络与应用集成点
- 5 华为香农实验室SDN-OF技术研发

# Overlay 与边缘Overlay网络

- Overlay网络是建立在另一个（物理）网络（IP)之上的一个具有“相似特性”的逻辑网络。
  - 从逻辑网络到物理网络的管理映射“全部”由自动程序实现
  - 通过网络的管理平面来处理服务器主机间的虚拟交换
  - （逻辑）链路可以通过建立基于底层网络的隧道实现
  - 能与任何硬件兼容
  - 可能是实现网络虚拟化的最佳途径
  - **WLAN, VPN:** 如物理网络一般，具有完备的 L2, L3 转发功能及 ACLs
- **边缘Overlay网络:** 存在于物理网络边缘节点之间的软件结构。
  - 边缘: 虚拟网络首次与应用交互，接收流量之处
  - 这些节点之间通过各种“隧道”实现



# 智能边缘Overlay网络



- 智能边缘: 在虚拟主机和现有物理网络之间的智能抽象层
  - 把一些通用智能型控制应用（如负载均衡，移动性，故障切换，分离和安全等）作为边缘上的Overlay
    - 它使与任何网络结构（fabric）设计的兼容性成为可能
    - 更重要的是, 边缘拥有大量丰富的有关真正端-端寻址，安全上下文，会话，移动性等语义信息。在边缘实现相关功能使得系统建造者可以持续改进而无需改变网络结构（fabric）
  - 中央控制器为虚拟边缘线卡编程配置转发策略，并支持任何潜在的上游云管理平台（CMP）如 OpenStack，VMware vCloud，CloudStack；配置虚拟网络，并支持任何潜在上游云管理平台（CMP）如 OpenStack API 接口VMware vCloud，CloudStack；提供全局视野并能够 随时随地按需要来实现程控
  - 虚拟网络的任何网络配置，只需要在边缘存在并由控制器自动编程
  - 管理虚拟的网络地址空间，并在边缘实时完成网络虚拟视图与物理视图之间的映射
  - 统一的L2-L7管控机制

- 智能化的网络边缘增加了网络智能信息，同时为在应用和底层网络基础设施之间进行更多的信息交换，以及管理应用的优化提供了基础
  - 通过应用来调用端到端路由计算模块来计算最佳路由
    - 控制器通过标准北向接口提供所需信息
    - 所计算出来的最佳路由再通过控制器下发到各相关网元上
  - 边缘控制的路由 – 一种新的域间协同机制
    - 可以充分使用某些现有网络架构或协议
    - 同时可以尽快获得SDN-OF集中管理技术带来的低复杂性，和成本
    - 可以基于各个自治域的质量，可靠性要求进行选择配置
    - 为充分发挥SDN-OF带来的好处，可能需要扩展现有OF协议，以有效与现有通用协议（MPLS,RSVP...）互通
  - 其他可能高价值应用：
    - 带宽保证
    - 流量整形控制（shaping）
    - 链路故障恢复



# 目录



- 1 SDN集中中央管理使智能控制优化成为现实
- 2 智能控制必须广义-支持异构网络/设备/协议
- 3 智能控制系统应该可编程，自适用，自管理
- 4 智能控制信息来自边缘-网络与应用集成点
- 5 华为香农实验室SDN-OF技术研发

# 聚焦关键技术



- 模式驱动架构和开发 (MDA-D)
- 广义控制器：一个系统同时支持OpenFlow 1.0/1.2/1.3 (以及传统网元)
- 自动发现全网拓扑，协议版本协商，能力学习（特别是对多流表及其相应Action的匹配）
- 中央多控制器集群实现对一个“较大”的自治域的全控，同时实现自动负载均衡和故障恢复 (OF1.2)
- 基于典型应用和优先级的多流表设计和有效实现 (OF1.2)
- 容错/快速恢复 (OF1.2)
- 用meter 表来实现流量管控 (OF1.3)
- 最小带宽保证 (OF1.3)
- 业务和应用策略驱动的路由算法
- 对控制器和控制通道恶意攻击的检测和防护
- 基于全网拓扑，状态和流量特征的多路径算法动态实现资源的有效利用，和快速建立保护路由
- 北向可以标准化的API:
  - 与云平台 (Openstack/Quantum)的集成
  - 与路径算法，网络智能分析，SDN-TE等的集成

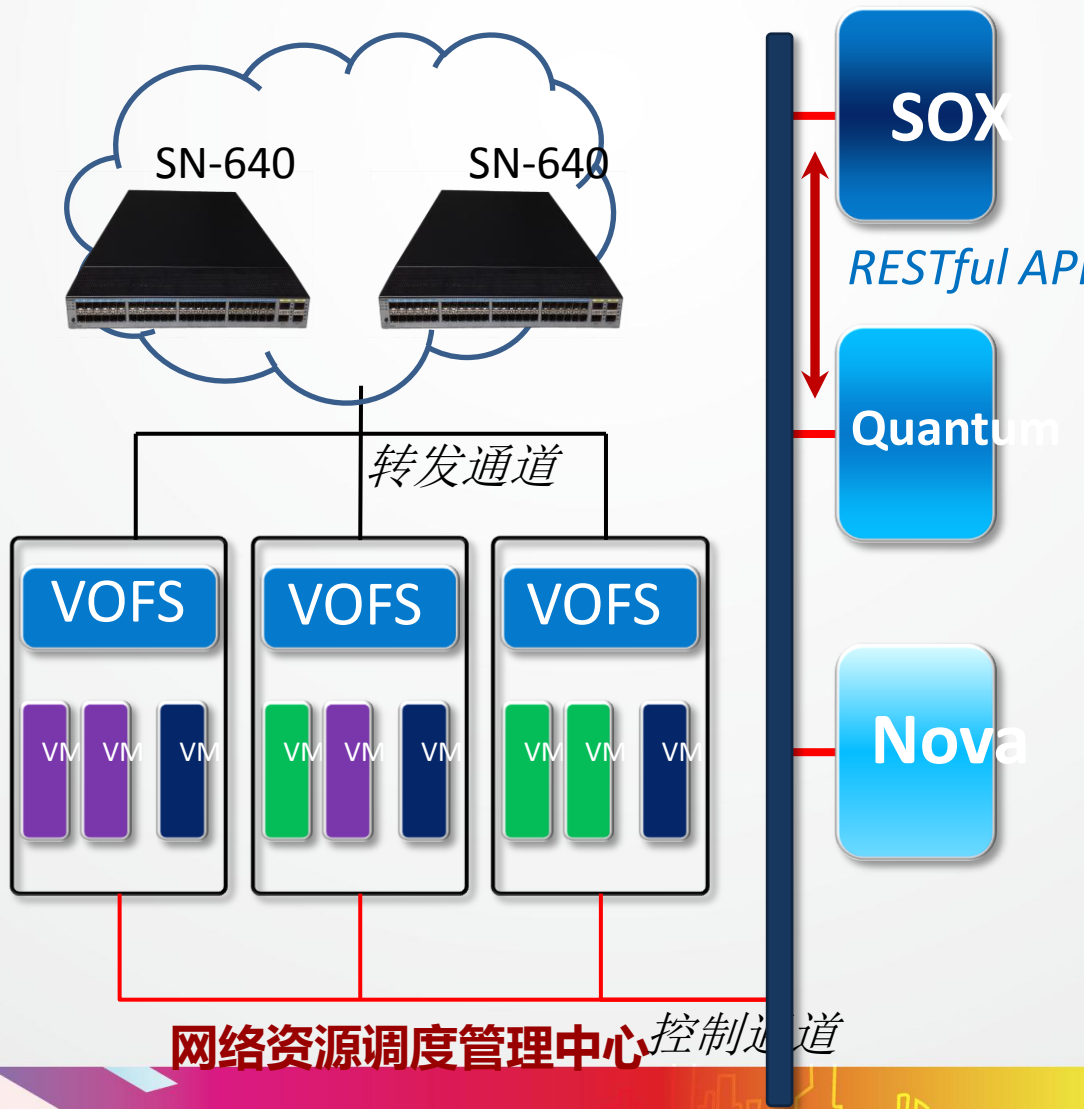
# 参加业界公开测评，确立业界技术领先地位



- 2012年10月ONF PlugFest
  - 唯一通过OF1.2测试厂家：MPLS,IPv6,。。。
  - 广义控制器对OF1.0和1.2混合网络的管控
  - 多控制器
  - 对典型网络应用优化的多流表设计和实现
- 2013年2月EANTC SDN HotStaging以及2013年3月SDN峰会
  - IPv6多播，MLD
  - 容错/快速恢复
  - 节点/链路故障自动检测和路由
  - 策略路由
  - 根据交换机能力所定的业务规则自动配置特殊路由
  - 完成不同交换机对标签需求不同的自动匹配和转换
  - 支持多网融合
- 2013年6月ONF PlugFest
  - 唯一完成OF1.3关键特性和应用的厂家：QoS流量管控，最小带宽保证
  - 完成对OF1.0,1.2测试床的所有测试
  - 完成由OF1.0/1.2/1.3交换机混合组网并成功管控

# SDN与云平台的完美集成

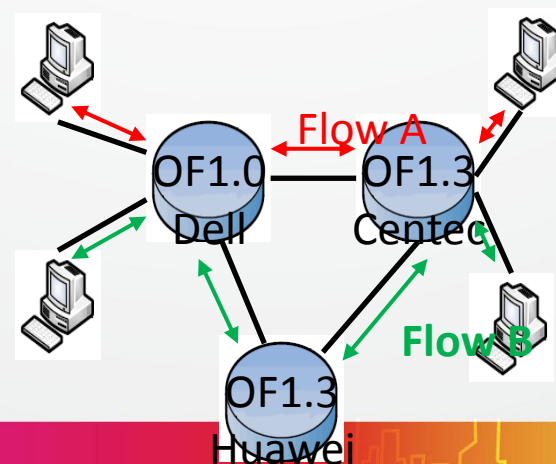
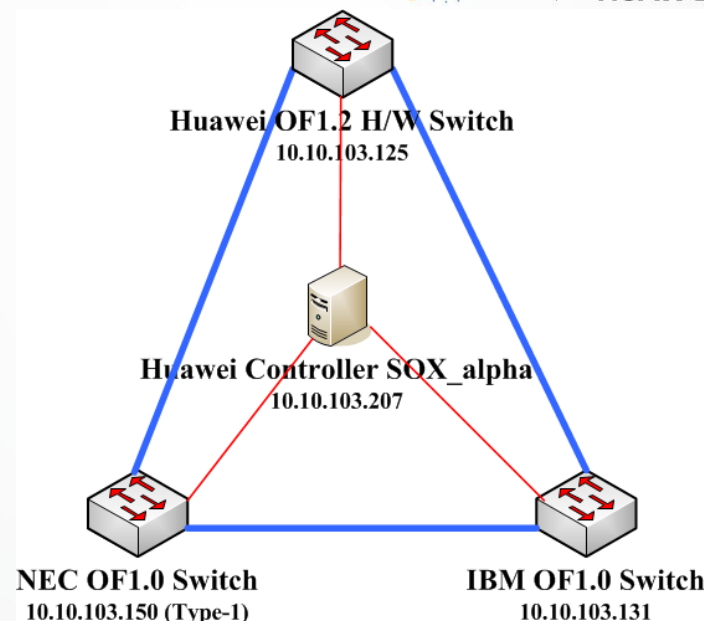
- 整合OpenStack云平台，使用SOX提供基于Openflow1.0至1.3的全版本网络可管可控
- 可支持多种虚拟化技术的软件交换机(VOFS)
- 无阻塞多流表转发架构的OpenFlow硬件ToR交换机(SN-640)
- 基于ROA的SDN北向API
- 最大化资源利用率的动态多路径技术和租户带宽保证技术
- 用户自定义的策略路由



# Demo 1 - 华为广义控制器技术 (OF1.0,1.2,1.3,...)

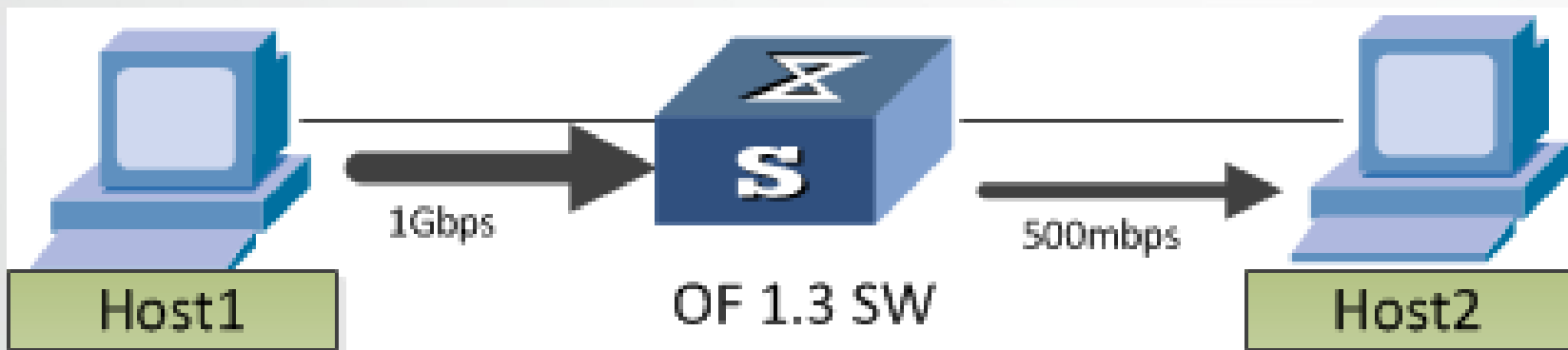


- 自动拓扑发现
- OpenFlow协议版本发现与协商
- 交换机能力学习，包括对多流表的支持
- 策略路由
- 不久（？）的将来：多网融合技术



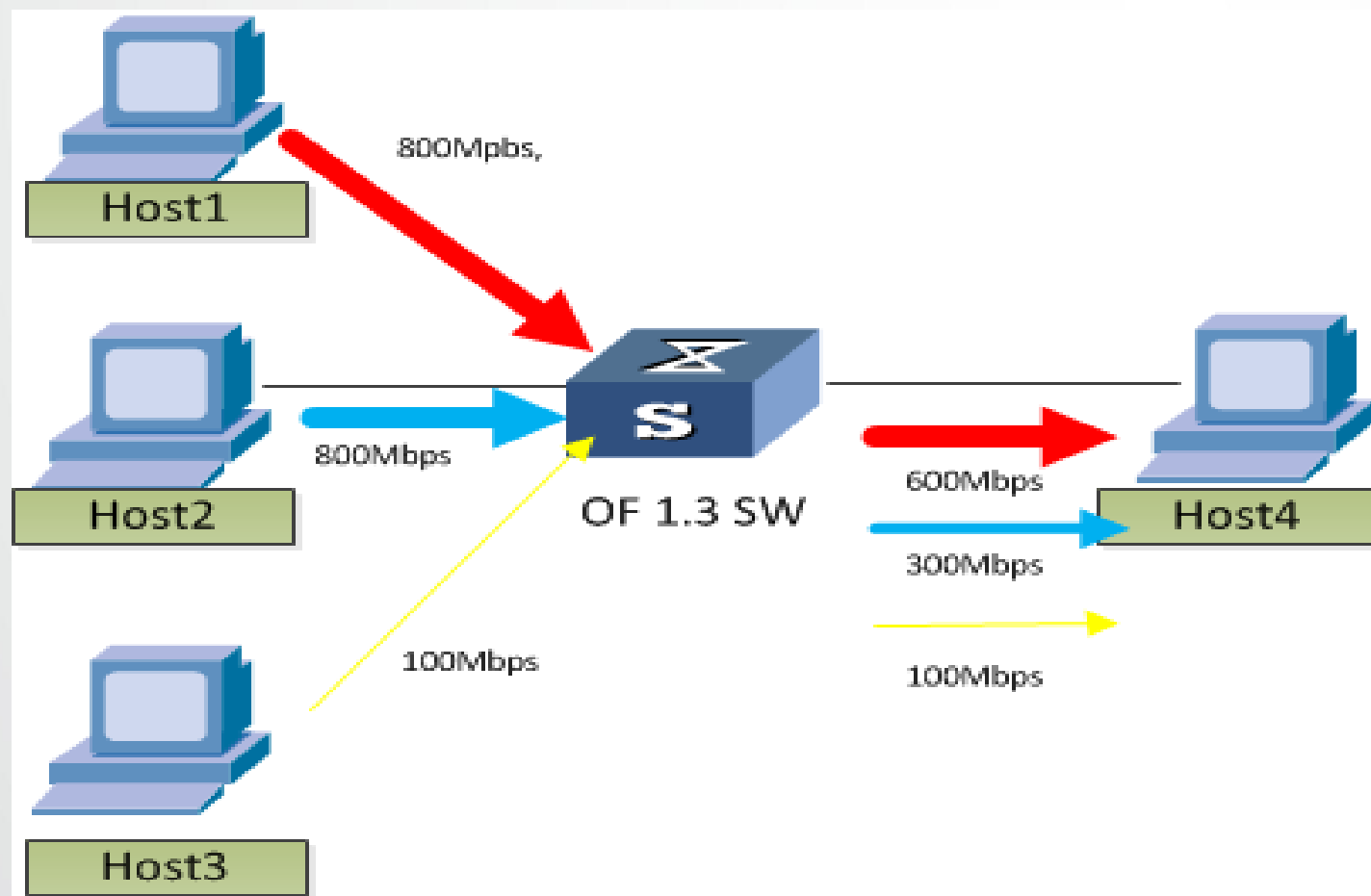
## Demo 2 – 用Meter表实现 流速限制（Rate-limiting）（OF1.3）

- OF1.3 引入了meter表来实现对每一个流的统计。统计的结果可以用来实现简单或复杂的QoS操作
  - 流速限制（Rate-limiting）

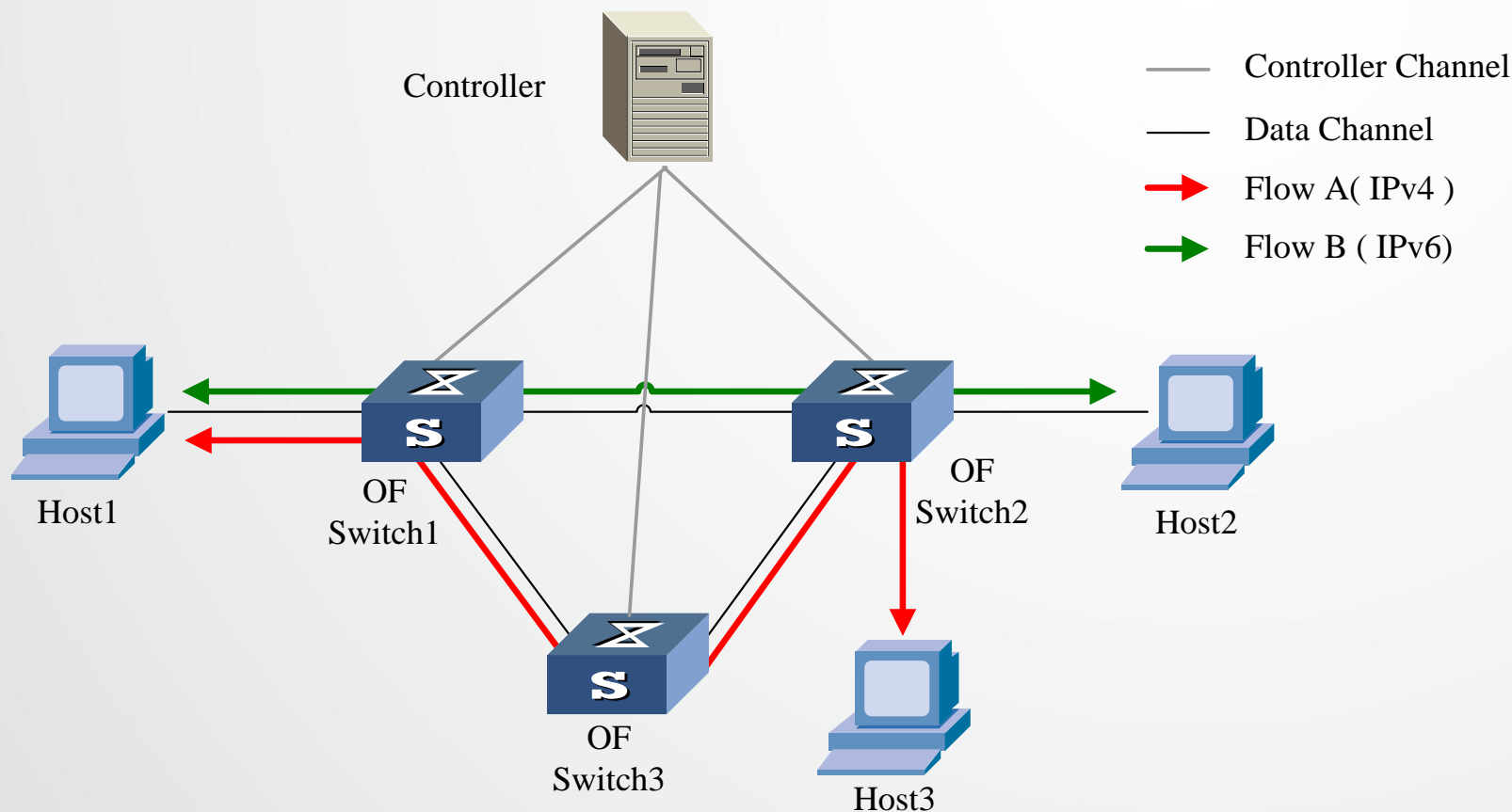




# Demo 3 – 用队列实现最小带宽保证 (OF1.3)



# Demo 4 – 在同一OF网络上 转发IPv4 和IPv6包





# Thank you

[www.huawei.com](http://www.huawei.com)

Copyright©2012 Huawei Technologies Co., Ltd. All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.