

DIM-B203

Windows Server 2016中 的软件定义网络

张力

Senior Program Manager
Microsoft Azure Stack

议程

为什么要软件定义网络

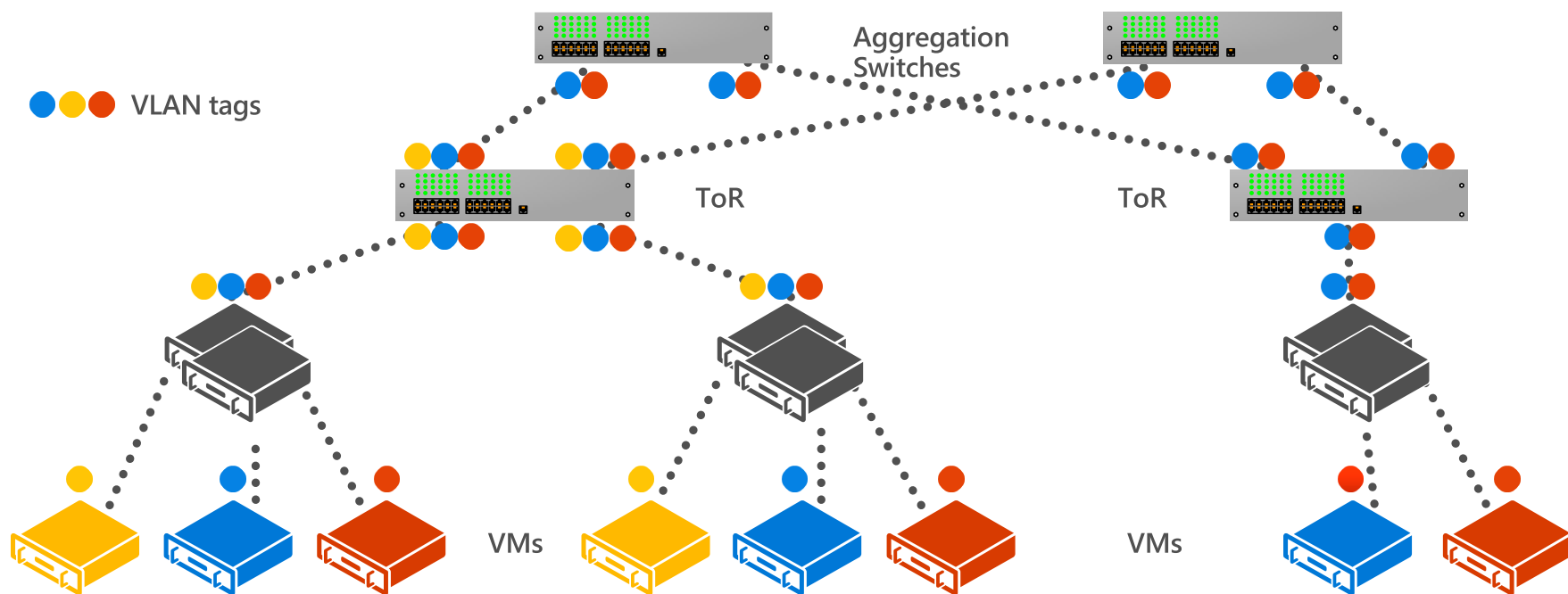
WS 2012 R2中实现的软件定义网络及其局限

WS 2016中的SDN基础架构

网络功能虚拟化 (NFV)

为云规模的部署而设计

传统的网络是如此不灵活

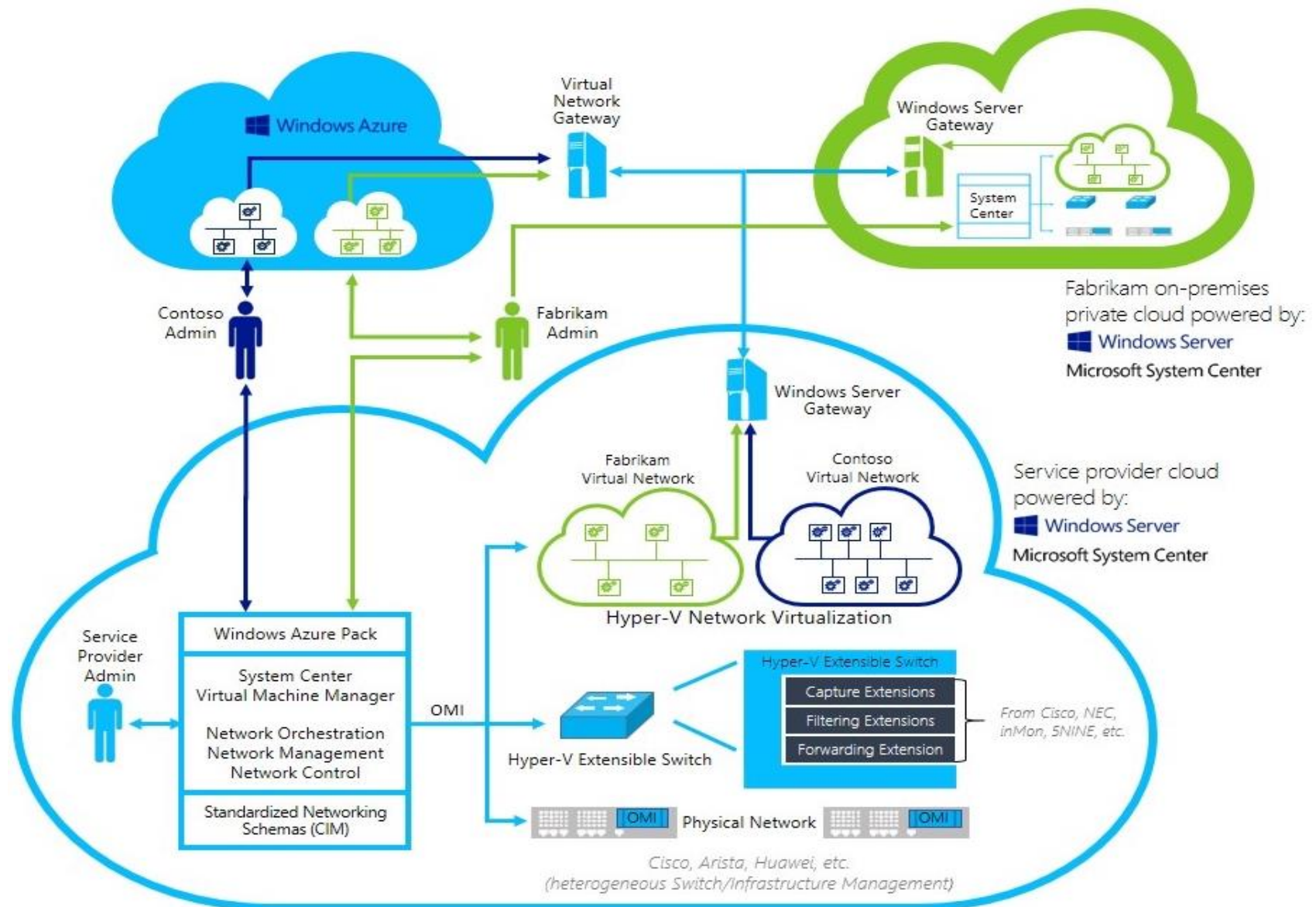


这样的拓扑限制了虚拟机的移动，需要重新配置生产环境的交换机

软件定义网络可以帮你：

- 简化物理网络拓扑
- 利用通用的交换设备和线缆
- 减少网络端口、交换机和核心交换机端口的数量
- 减少专用的设备
- 将原本网络设备的功能交由虚拟机和软件来完成
- 通过软件来进行优化
- 通过软件来实现隔离

WS 2012 R2中的软件定义网络



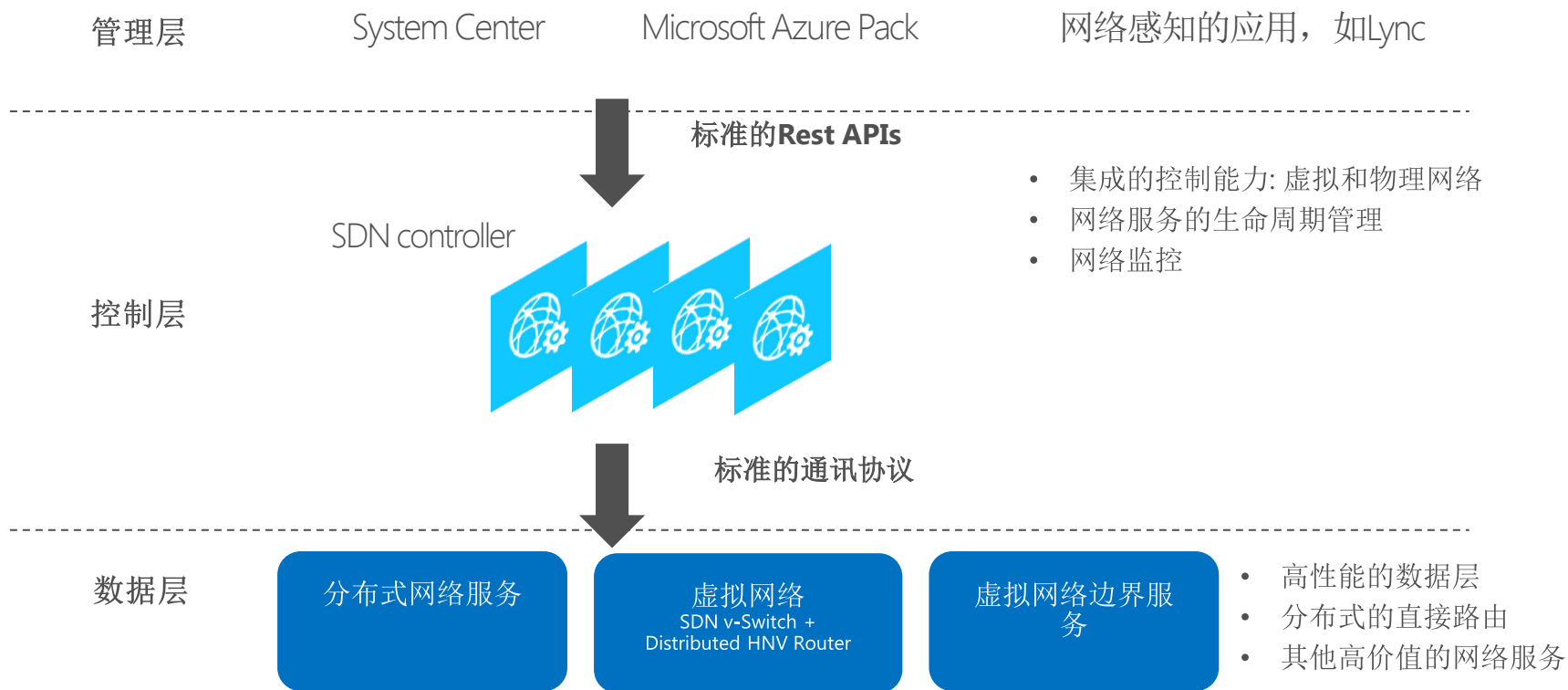
2012 R2中的局限

- 没有内置软件负载均衡器
- 没有多租户的直接路由网关
- 给一个租户虚拟网络分配多个外部IP地址
- 没有访问权限控制（ACL）

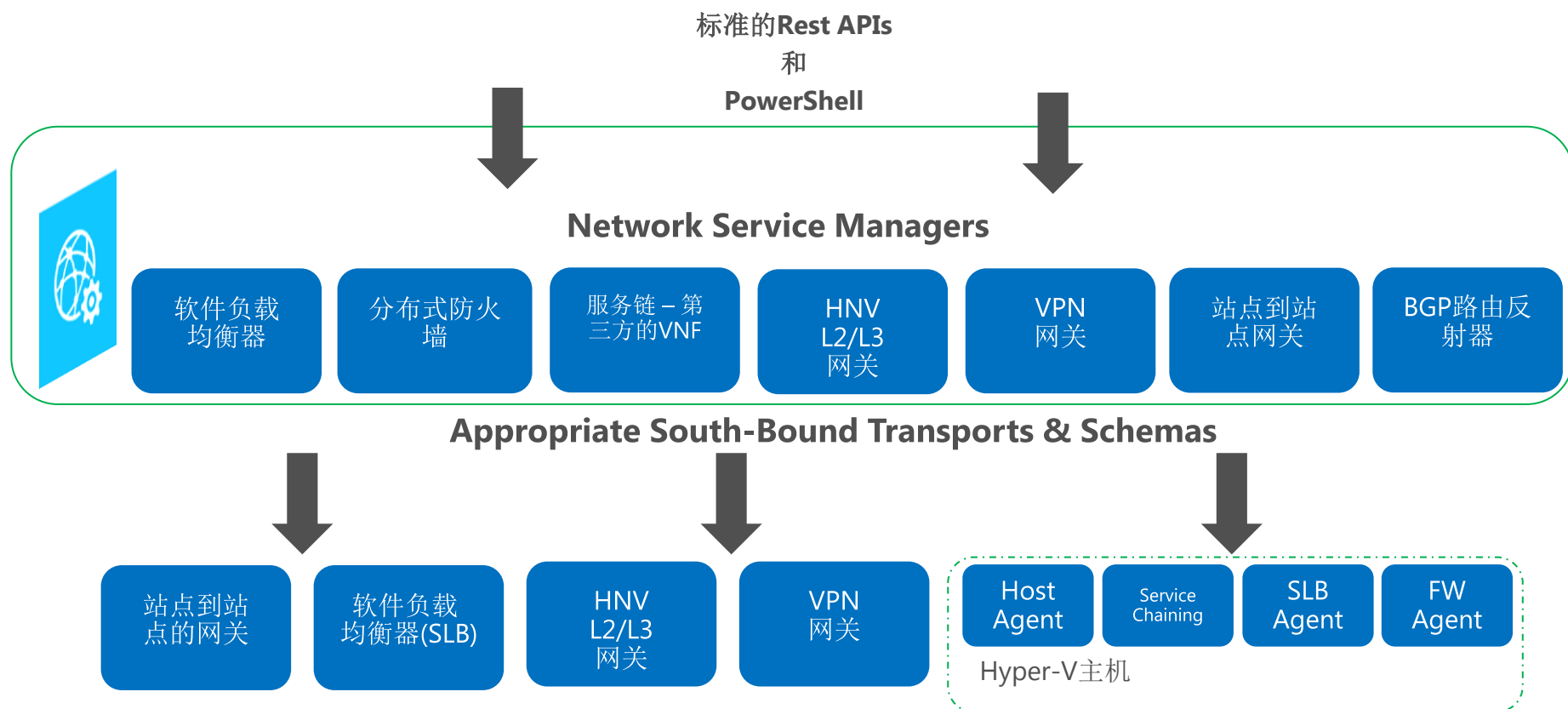
WS2016在网络方面的重点投入

SDN基础架构	网络功能虚拟化(NFV)	为云规模的部署设计
可横向扩展的Network Controller，可以通过服务器或者虚拟机来实现	可以由租户自己来定义服务链条（包括经由哪些NFV设备，如何路由，作何控制等）	融合底层架构支持将租户的虚拟网络流量和使用RDMA的存储IO融合在一起并提供相应的QoS
增强在每一层的和其他友商产品的互操作性	内置在Azure中已经被广泛应用的成熟虚拟网络服务	提供高吞吐、低延迟的数据包处理架构
借助SCVMM集成部署，生命周期管理，自动化等能力来简化管理； 同时由Azure Stack提供租户自服务能力	对于基于Linux的第三方虚拟网络设备提供最佳的性能	与业界领先的厂商合作提供高质量的经过优化的Offloading处理
		源自公有云的DNS服务（包括能够感知请求来自的地域，流量管理和增强的IP地址管理）

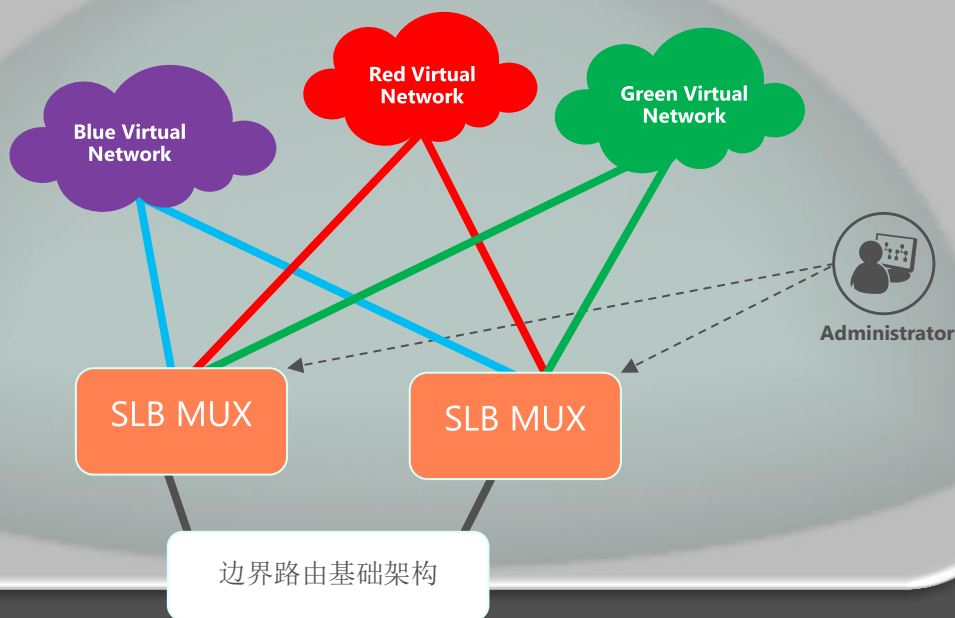
SDN基础架构



网络功能虚拟化 (NFV)

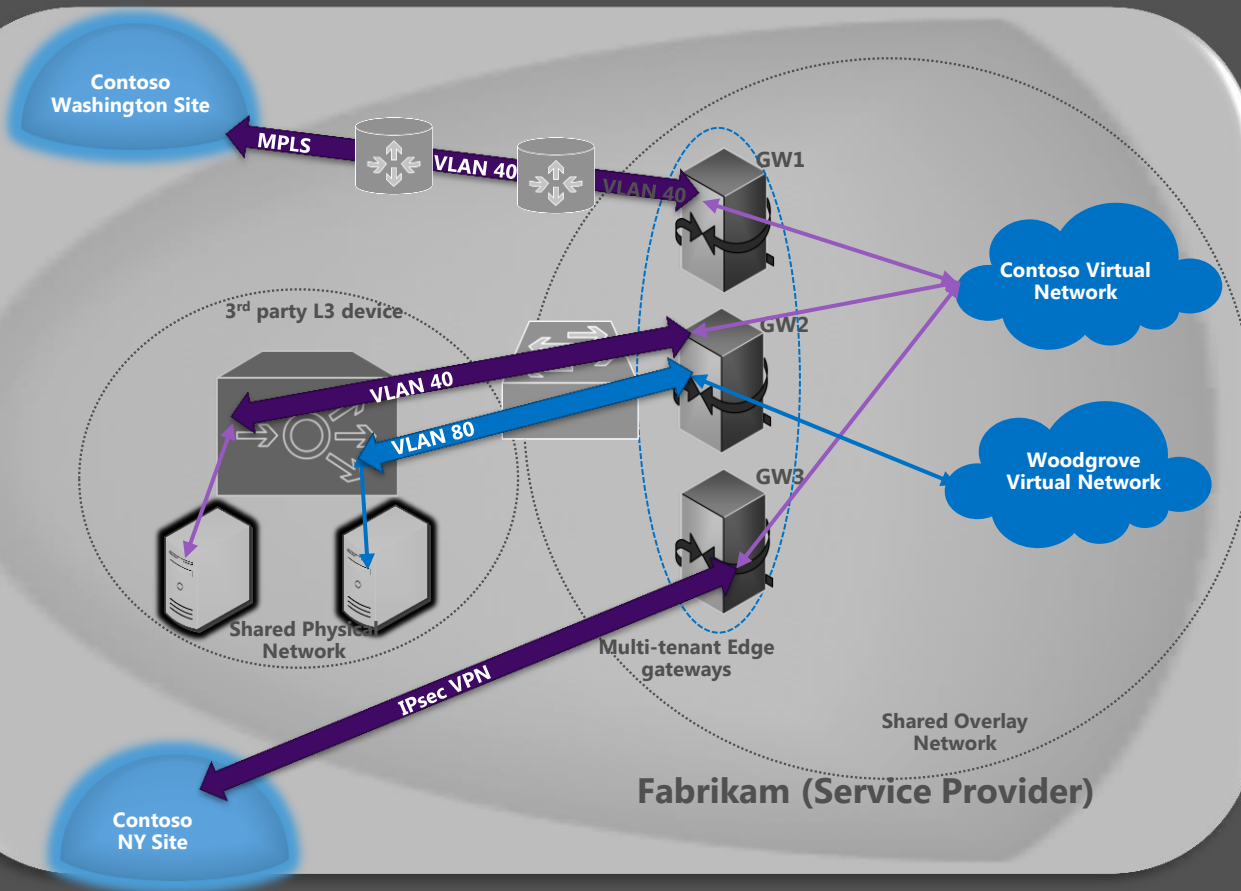


软件负载均衡器 (SLB)



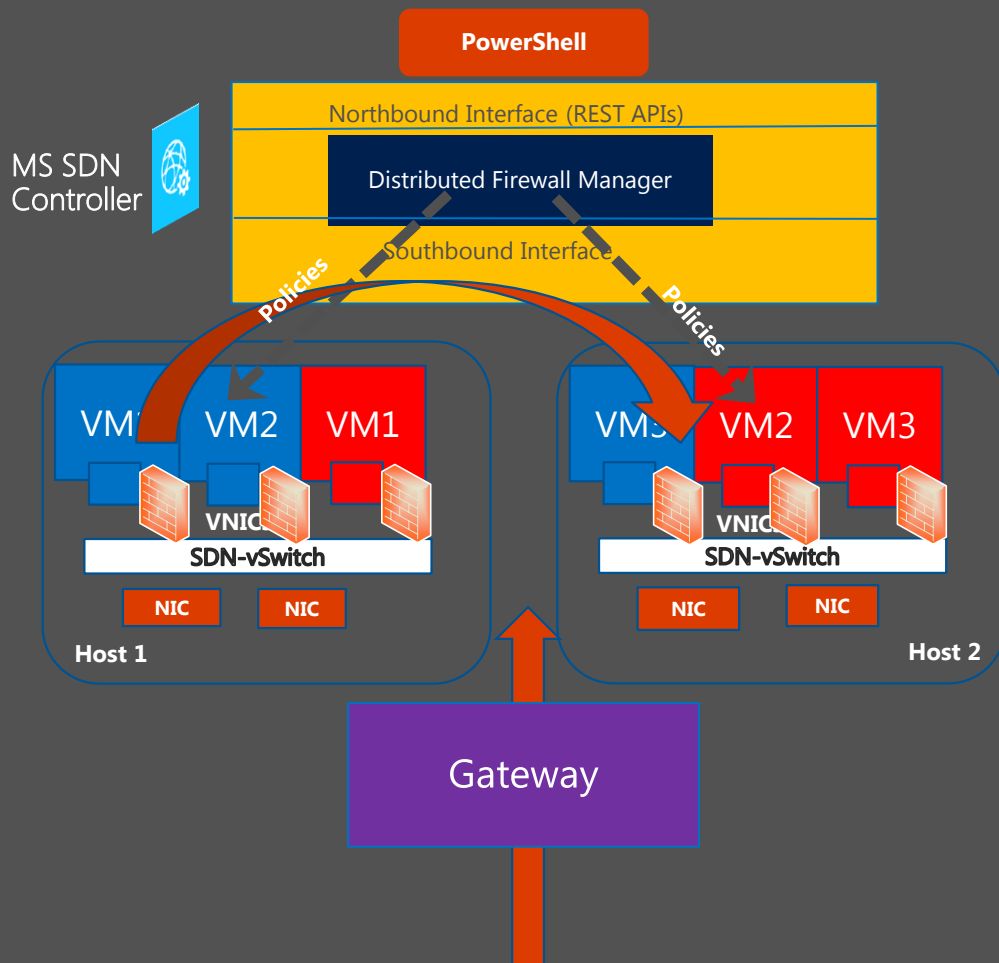
- 多租户：四层软件负载均衡服务
- 高可用：通过BGP路由来实现动态的冗余
- 纵向扩展：到10 Gigabit以太网

多租户的HNV直接路由网关



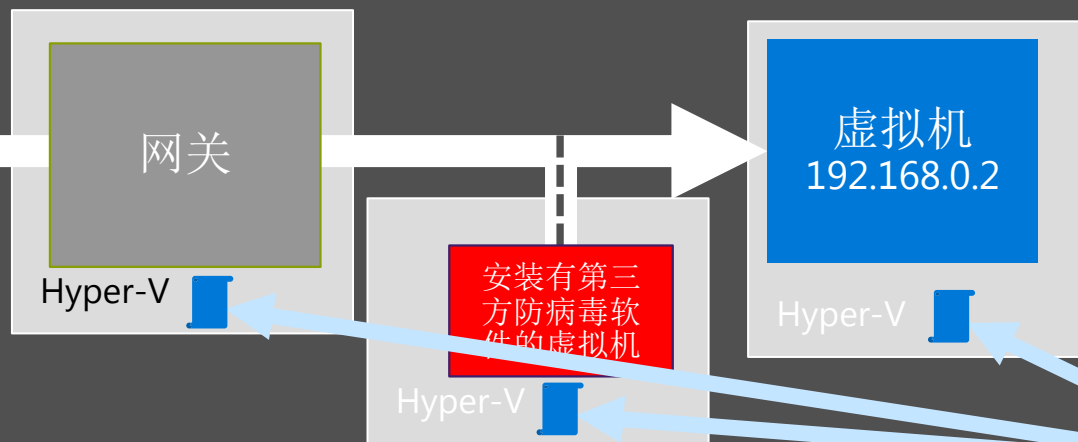
- 集成HNV租户网络和使用VLAN隔离的三层网络
- 单个网关支持多个租户的VLAN
- 静态路由或者BGP
- 中间的交换机或路由器需要支持基于VLAN的隔离
- 部署和高可用同S2S网关一样 (M:N)

分布式数据中心防火墙



- 分布式的多租户防火墙保护虚拟网络
- 保护策略作用在每一个租户虚拟机的SDN-vSwitch的端口上
- 保护所有的流量，包括东西和南北
- 同Network Controller集成
- 保护策略可以通过租户门户递入，而后Network Controller将他们分发到所有适用的物理主机上

服务链条



Rule

SourceIP=Any
DestinationIP=192.168.0.0/24
Protocol=Any
SourcePort=Any
DestinationPort=Any

+

Service Chain

Element1="3rd Party Antivirus VM"

+

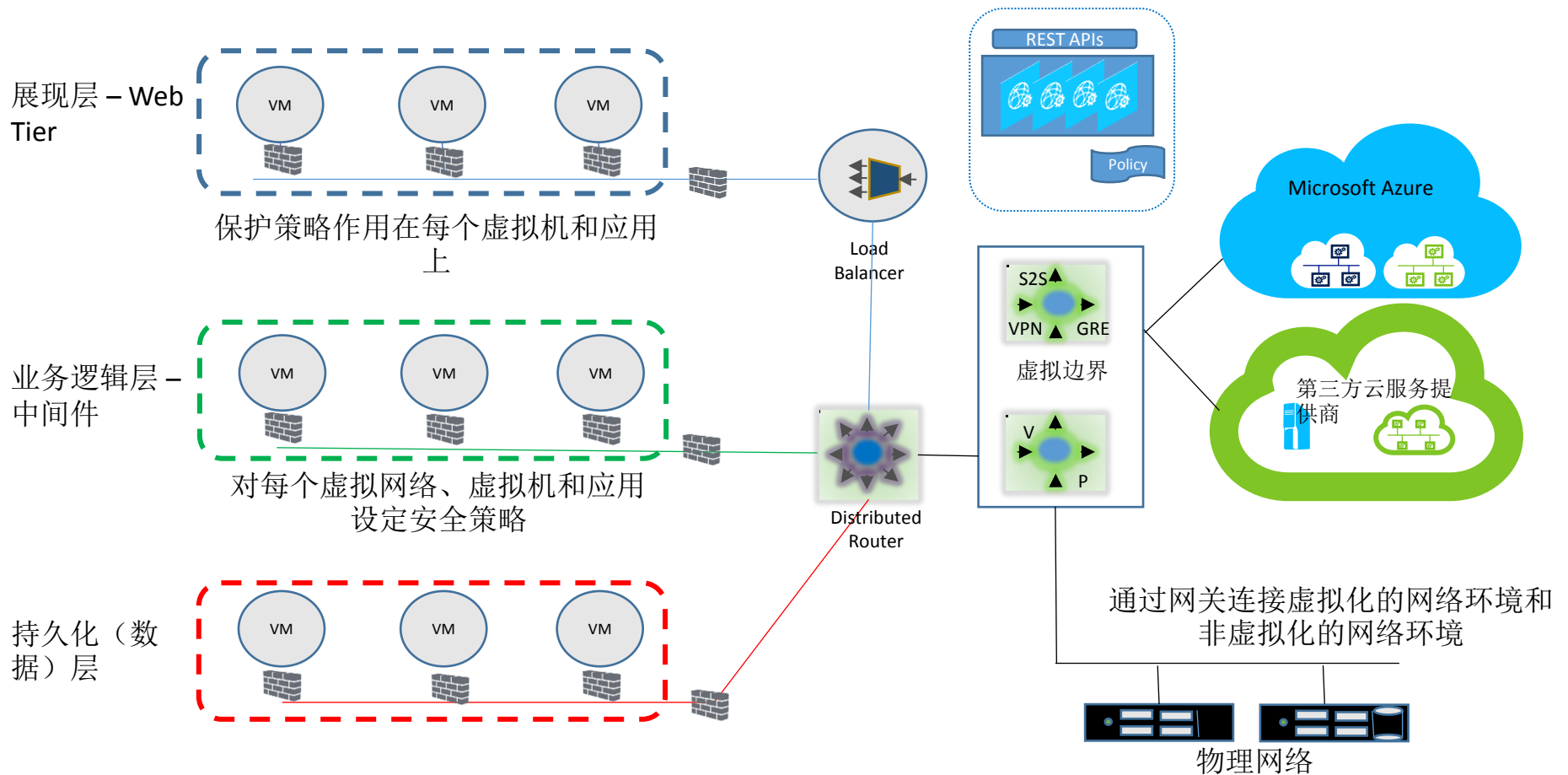
Group

Virtual Network="MyNetwork"



Network Controller

传统的多层应用部署环境中运用SDN提供的能力



SDN演示

为云规模的部署设计

融合的架构	Packet Direct	集成Offload管理
单块网卡可以同时提供启用RDMA的对数据存储的访问和支持租户的流量	10GB -----> 40GB/100GB	同合作伙伴一起关注质量从而提供可靠的，可以预期的云规模的部署
集成了SDN v-Switch和NIC teaming	微秒级的延迟	优化的 offload使用: checksum offloads, LSO, RSC, RSS, VMQ, SR-IOV, NVGRE TO, Routable RDMA, VXLAN TO, PD integration
集成了QoS – 在多个不同类型的流量中合理的分享带宽	每个CPU核心支持几百万的PPS	

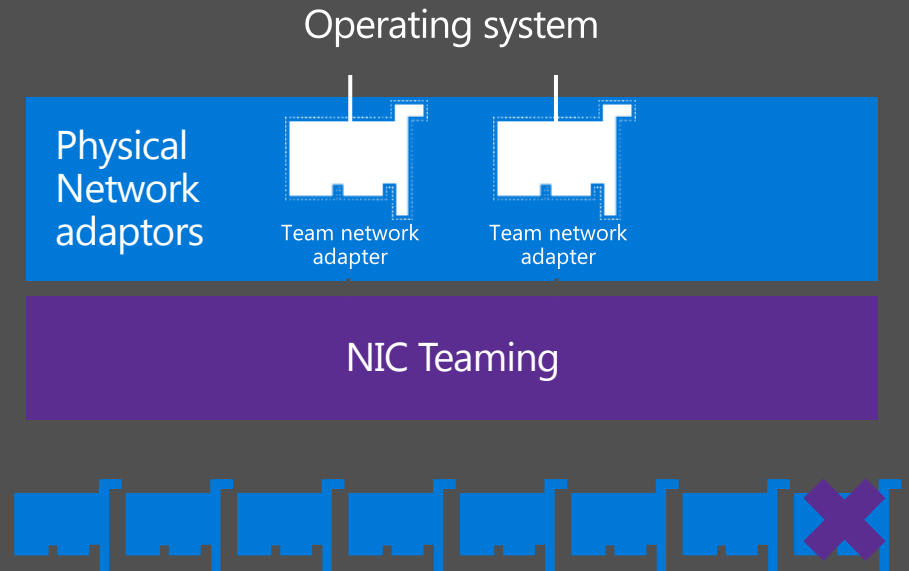
如何进一步提高性能

- 在标准的网卡硬件上实现更多的带宽，更多的PPS，更低的CPU占用率
- PacketDirect是2016新引入的功能，它能够让网络性能有质的飞跃

与交换机无关的网卡Teaming

网卡可恢复性和负载均衡的集成结局方案

- 与供应商无关并且内置
- 最多可以组合32张网卡
- 聚合多块网卡的带宽同时还提供在某块网卡故障时的流量的切换
- 包括多种模式：基于交换机的和独立于交换机的
- 多种流量分布算法：Hyper-V Switch Port, Hashing和Dynamic Load Balancing



RDMA的考虑

■ 为什么要RDMA

- 大大降低在Hyper-V计算节点和横向扩展文件服务器中CPU的开销
- 允许将更多的计算时间留给运行的虚拟机
- 改善数据传输速度

■ 两种主要的基于以太网的选择

- RoCE：Connect 3 Pro也可以路由，需要交换机支持PFC，需要安装DCB
- iWARP：可以路由

■ 那种更好？应考虑：

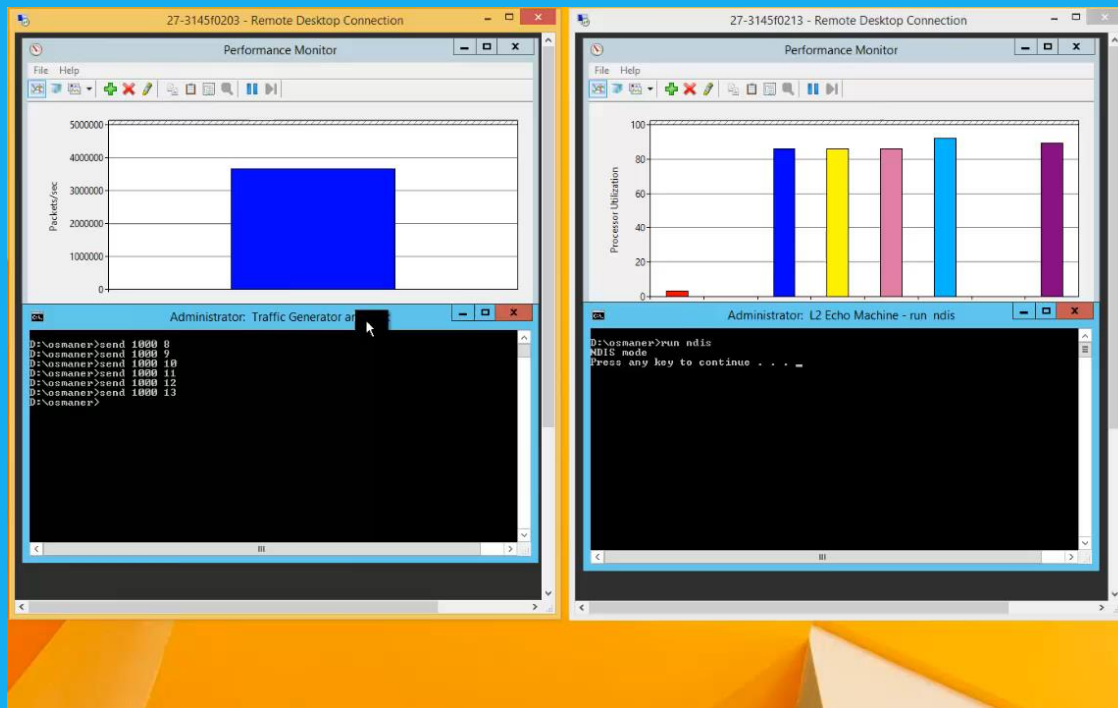
- 吞吐量
- 安装复杂性
- 部署环境的需求，比如是否需要路由？

■ Infiniband也是一个很好的选择如果你已经有IB

RDMA演示

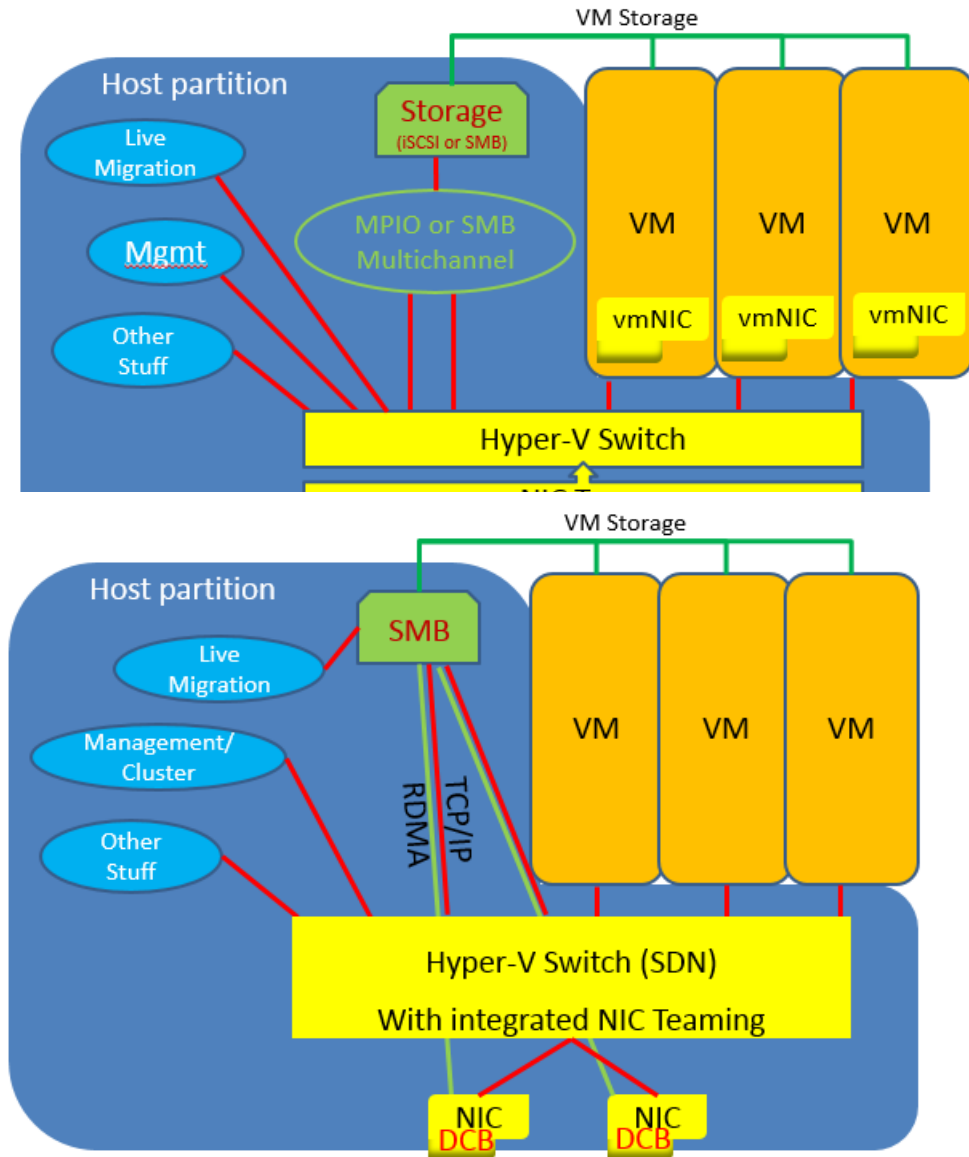
	SMB DIRECT	SMB
IOPS	195,060	203,149
IOPS STDEV	4286	39640
LAT	1.312	1.259
LAT STDDEV	0.754	117.61
CPU USAGE	10.16%	16.42%

PacketDirect演示



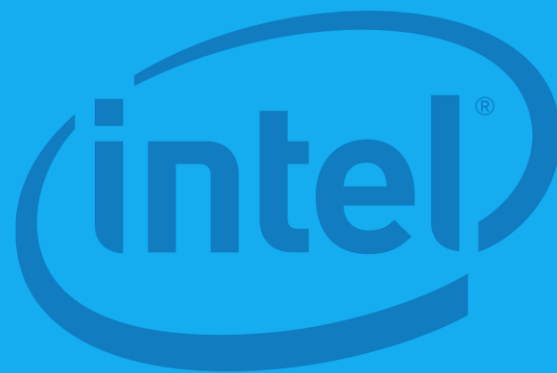
融合网卡

- 2012 R2上，RDMA需要独立的网卡，从而和以太网的流量分开且不能做 Teaming
- 2016中，融合网卡成为可能（融合访问的存储的 RDMA流量和以太网的流量）
- 通过QoS在两种流量中进行隔离



融合网卡的演示

感谢！



来自ECG产品研发团队的其他讲座

10/27

10:00AM -10:50AM	DIM-B305: 在Hyper-V中保护租户的秘密
1:30PM – 2:20PM	DIM-B201: Azure Stack: 与Azure一致的私有云、托管云服务交付
3:30PM – 4:30PM	数据中心专家面对面
4:30PM – 5:20PM	联想-微软Cloud Rack实践及Window 2016 私有云存储

10/28

9:00AM – 9:50AM	DIM-B306: 在私有云上构建和Azure一致的现代应用
2:30PM – 3:20PM	DIM-B204: Windows Server 2016中的软件定义存储
3:30PM – 4:20PM	DIM-B203: Windows Server 2016中的软件定义网络
10:00AM -10:50AM	动手实验室: DIM-H21 Windows Server 2016 中的软件定义存储
3:30PM - 4:20PM	动手实验室: DIM-H26 在 Hyper-V 中保护租户的秘密

关注Azure Stack, 加入我们线上线下的讨论

微信群

Azure Stack Friends

Yammer Group

https://www.yammer.com/wsscengineering/#/threads/inGroup?type=in_group&feedId=6394023

MeetUp

<http://www.meetup.com/Shanghai-Cloud-Service-Provider-Meetup/>



分会场课程视频观看及课件下载

大会结束后一个月内，敬请访问微软 Channel 9 官网观看更多课程视频及下载课件

<https://channel9.msdn.com/Events/Ignite/Microsoft-Ignite-China-2015>

或

<http://aka.ms/IgniteChina2015>



Channel 9

请您填写分会场课程反馈表

您的反馈将帮助我们改进

- 填写表格后，请离场时交给本会场的现场工作人员
- 谢谢！

您对本节课程的满意度：

(每行只勾选一项，请注意：9分代表非常满意，分数越少，满意度则越小，1分代表非常不满意)

本节课程整体满意度评价	9	8	7	6	5	4	3	2	1
您对讲师的专业水平是否满意？									
您对讲师的表达技巧与能力是否满意？									
您对 Demo 的解说及效果是否满意？ (如本节课没有 Demo 请留白)									
您对课程内容的安排是否满意？									

如您还有其他意见与建议：

谢谢！

