

Huawei Enterprise **A Better Way**

# 数据中心网络解决方案之 虚拟感知

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

HUAWEI TECHNOLOGIES CO., LTD.



# 目录

- **服务器虚拟化对数据中心网络的挑战**
- **华为虚拟感知解决方案技术原理**
- **华为虚拟感知解决方案应用部署**

# 云计算时代服务器虚拟化规模应用

- 69% of workloads virtualized by 2013<sup>1</sup>
- VM density rising to 8.4 in 2013<sup>2</sup>
- 53% VMware growth over last three years<sup>3</sup>
- 76% of enterprises pursuing private cloud strategy by 2012<sup>4</sup>

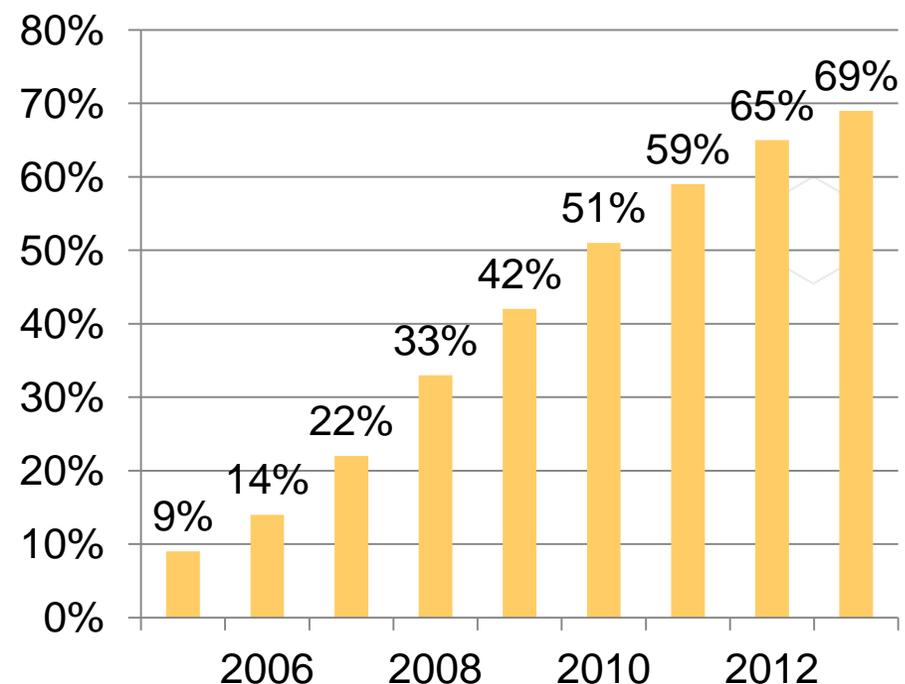
1. IDC, 2011, Market Analysis Perspective: Worldwide Datacenter Trends and Strategies.

2. IDC, WW Server Virtualization Shipment Forecast, 2005–2013.

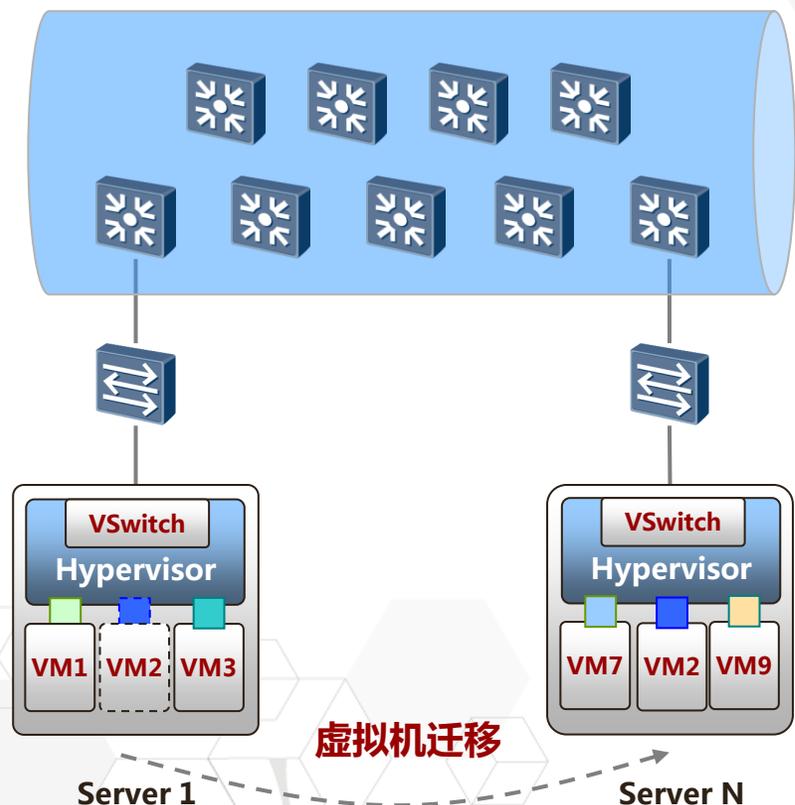
3. VMware published financial reports.

4. Thomas Bittman, Gartner Data Center Conference 2010.

## WW installed workloads virtualized



# 服务器虚拟化对数据中心网络的诉求（运维）



## ■ 管理职责不清

- VSWITCH在服务器内，属于系统管理员范畴；VSWITCH是交换机，又属于网络管理员范畴，VSWITCH谁来管理？
- 对VM的网络配置系统管理员和网络管理员如何协同工作？

## ■ VM故障定位困难

- VM的流动性，位置跟具体接入交换机无关。
- VM故障还是网络故障，服务器内VM交换流量网络管理员无法获知。

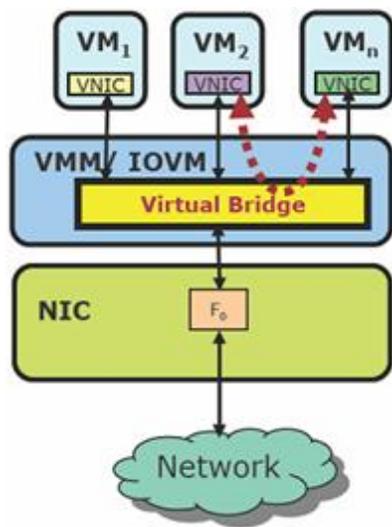
## ■ 网络配置动态迁移

- 网络配置如何随VM虚拟机的迁移而动态迁移？

## ■ VM的复杂策略控制

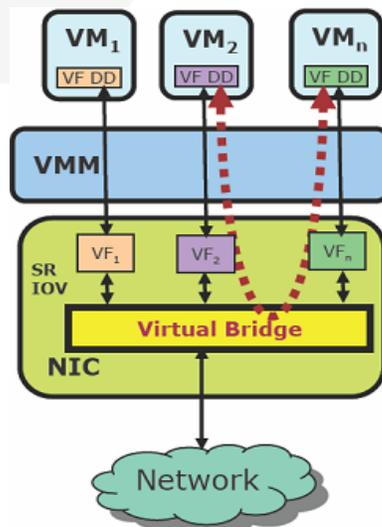
- VSWITCH交换机缺乏相应策略控制功能。
- 如通过VSWITCH交换机实现，需要耗费大量的CPU资源。

# 服务器虚拟化对数据中心网络的诉求（转发）



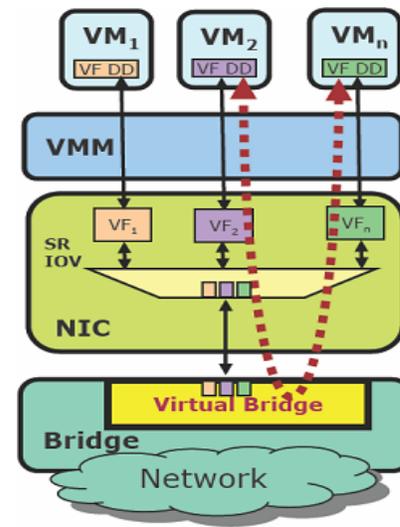
由软件交换机实现转发

优点	成熟可用
	TOR为通用交换机
缺点	占用服务器CPU资源



智能网卡实现交换机功能

优点	性能高
	TOR为通用交换机
缺点	网卡管理，互通性问题



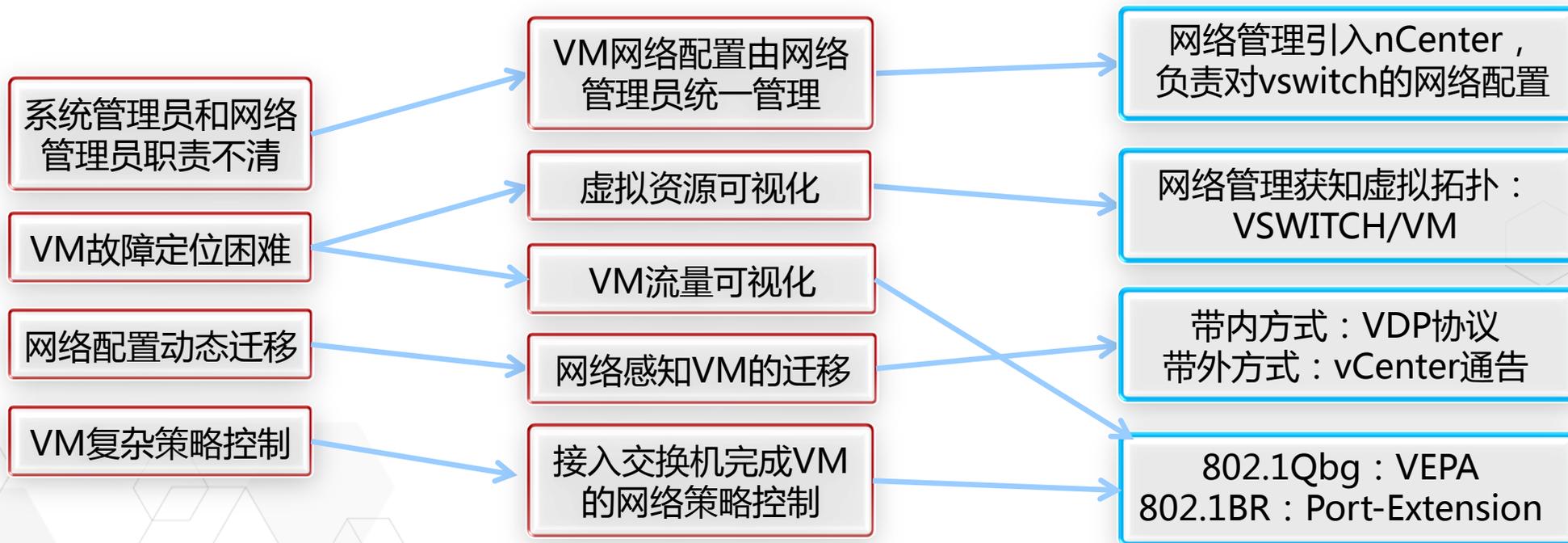
接入交换机完成环回转发

优点	性能高
	统一管理
缺点	TOR硬件支持转发模型

# 目录

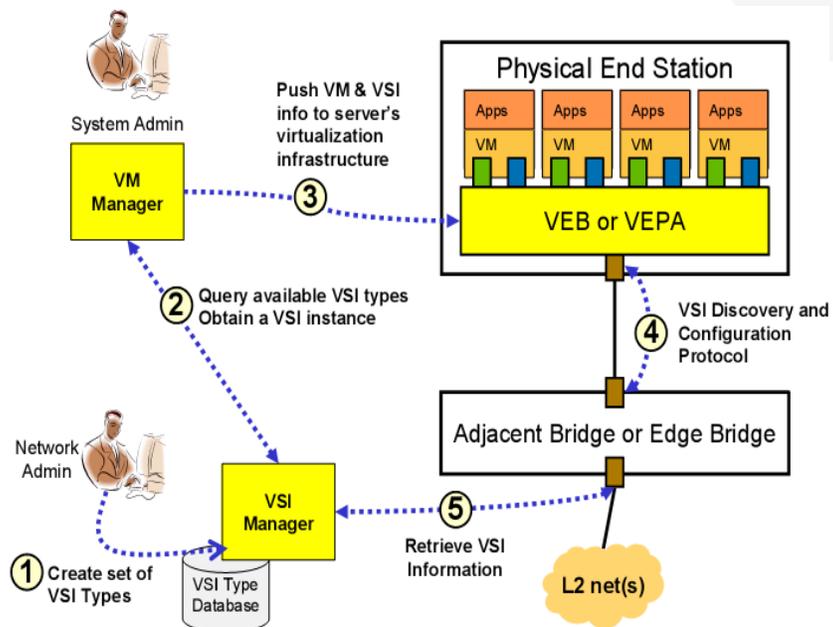
- 服务器虚拟化对数据中心网络的挑战
- 华为虚拟感知解决方案技术原理
- 华为虚拟感知解决方案应用部署

# 服务器虚拟化问题解决之道



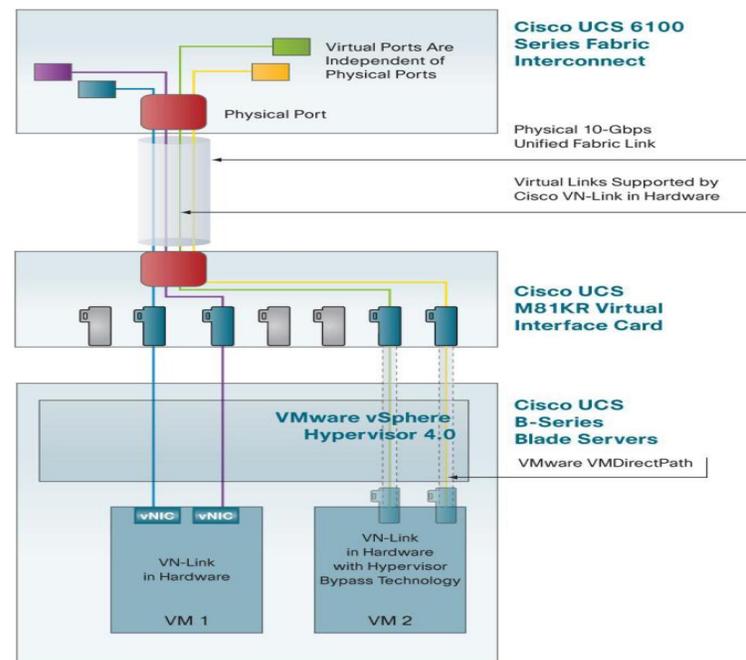
# 服务器虚拟化相关的数据中心网络标准

## 802.1Qbg



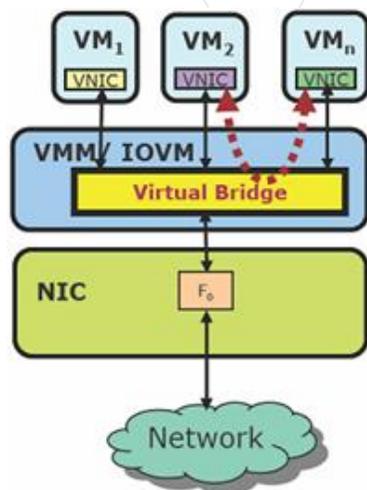
定义了服务器虚拟化相关的转发、发现、配置的处理机制。HP、IBM等众多厂商参与制订。

## 802.1br

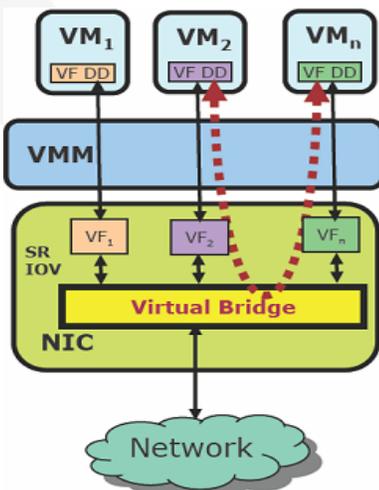


定义了虚拟端口通过外部交换机集中转发、集中控制的机制。由Cisco主推，原来是802.1Qbh。

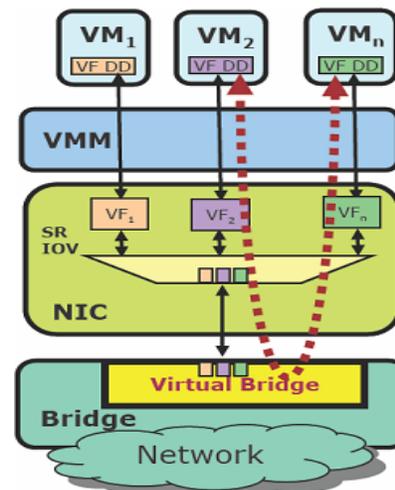
# 转发控制方案选择



由软件交换机实现转发  
802.1Qbg中定义为VEB



智能网卡实现交换机功能  
802.1Qbg中定义为VEB

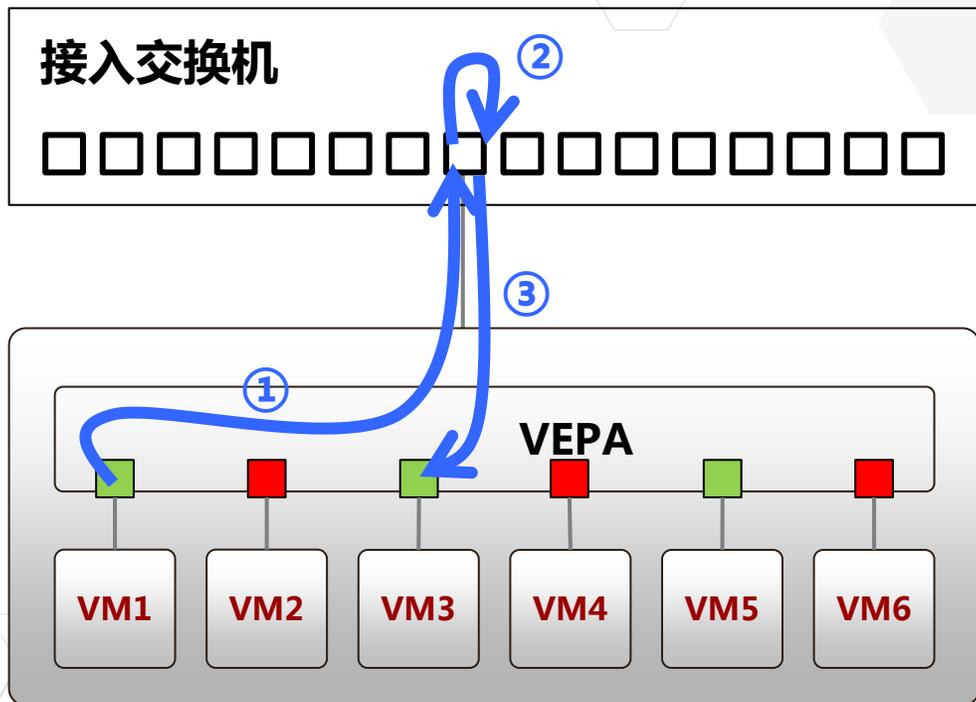


接入交换机完成环回转发  
802.1Qbg中定义为VEPA

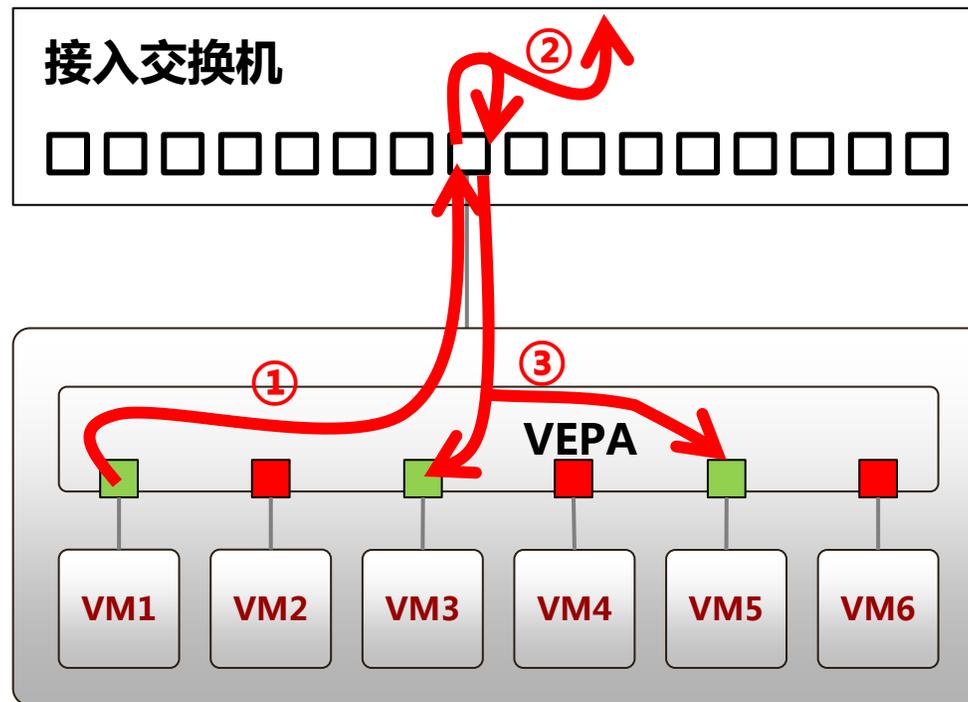
<b>VM流量可视化</b>	要实现流量采集，需要在物理服务器中创建一个虚拟机，专门用来运行探针	
<b>网络策略控制</b>	流量控制、ACL实现困难	流量控制由硬件支持，ACL数量少
	在接入交换机上完成流量采集	
	流量控制和ACL由接入交换机实现	

**华为虚拟感知解决方案支持三种转发模型  
推荐采用接入交换机环回转发方案满足管控要求**

# 接入交换机环回转发

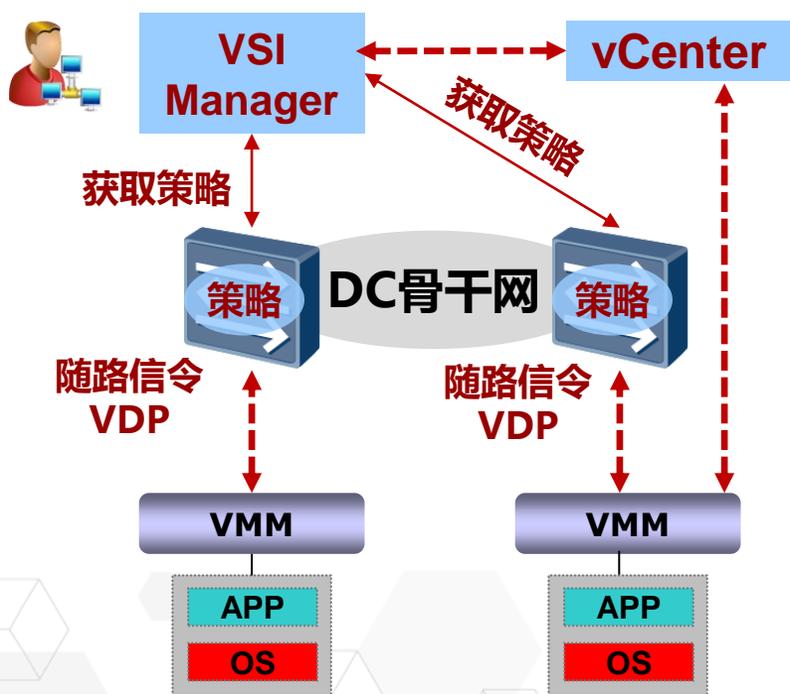


- VM1、VM3属于同一个VLAN
- VM1发给VM3的流量通过接入交换机环回回来
- 安全策略在接入交换机部署，功能丰富，不影响转发性能



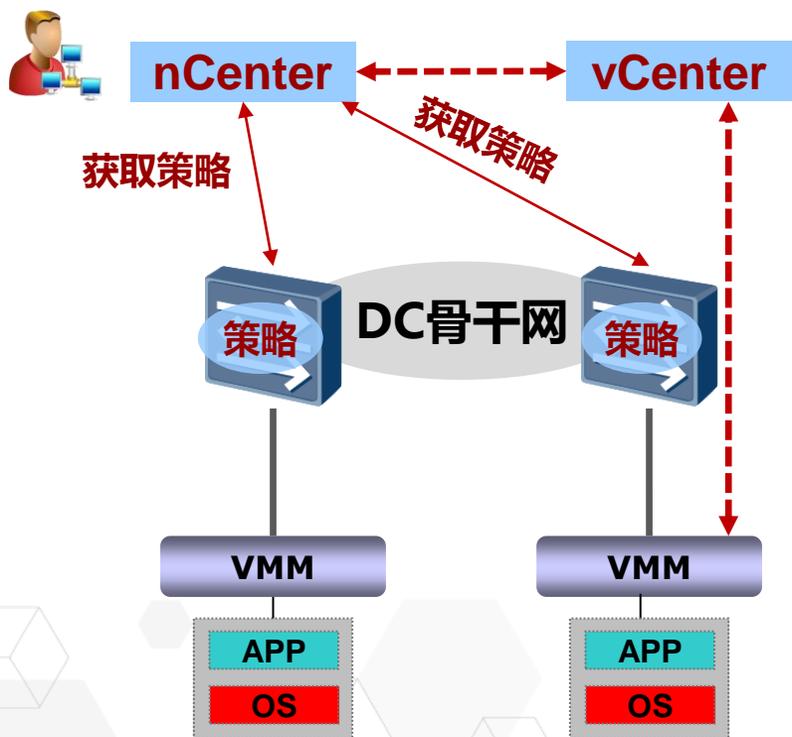
- VM1、VM3、VM5属于同一个VLAN
- VM1发出的广播报文，接入交换机向VLAN内其他端口复制，同时向虚拟交换机复制一份
- 虚拟交换机完成该VLAN的其他VM的复制

# 虚拟感知解决方案选择：带内管理



- 管理员在VSI Manager创建特定VSI类型的网络策略
  - vCenter从VSI Manager查询VSI类型和策略信息
  - vCenter将VM和VSI信息设置到服务器虚拟交换机
  - 服务器虚拟交换机通过802.1Qbg定义的VDP协议，将VSI类型信息通告给接入交换机
  - 接入交换机从VSI Manager查询特定VSI类型的网络策略，完成接入交换机上的网络策略配置
- 
- 由于虚拟机的发现是服务器虚拟交换机通过VDP协议通告给接入交换机的，这种方案称为带内管理方案
  - 802.1Qbg目前还没有成为正式标准
  - 业界虚拟机平台、交换机都尚未支持VDP协议
  - 802.1Qbg标准未具体规定和vCenter的接口，VSI Manager需要针对各虚拟机平台进行适配

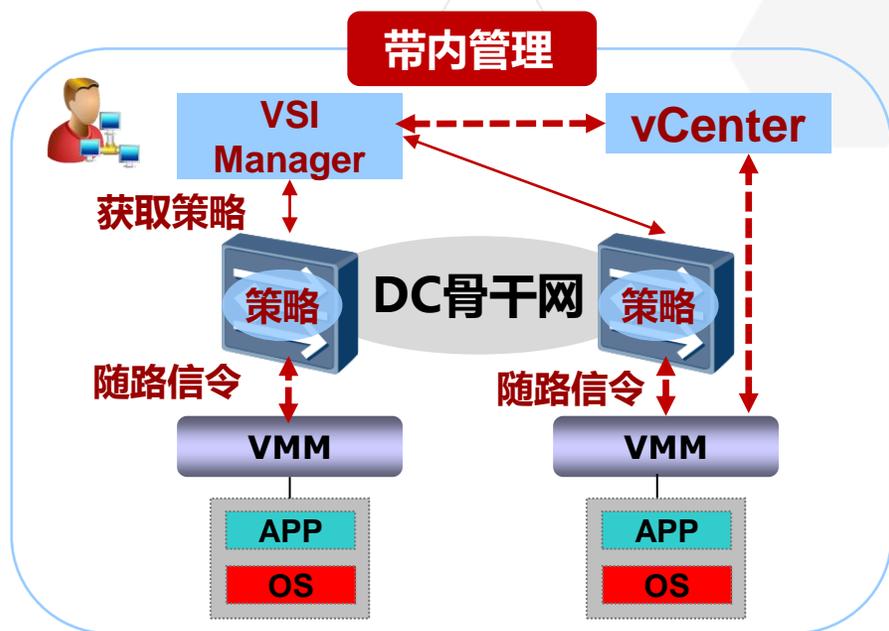
# 虚拟感知解决方案选择：带外管理



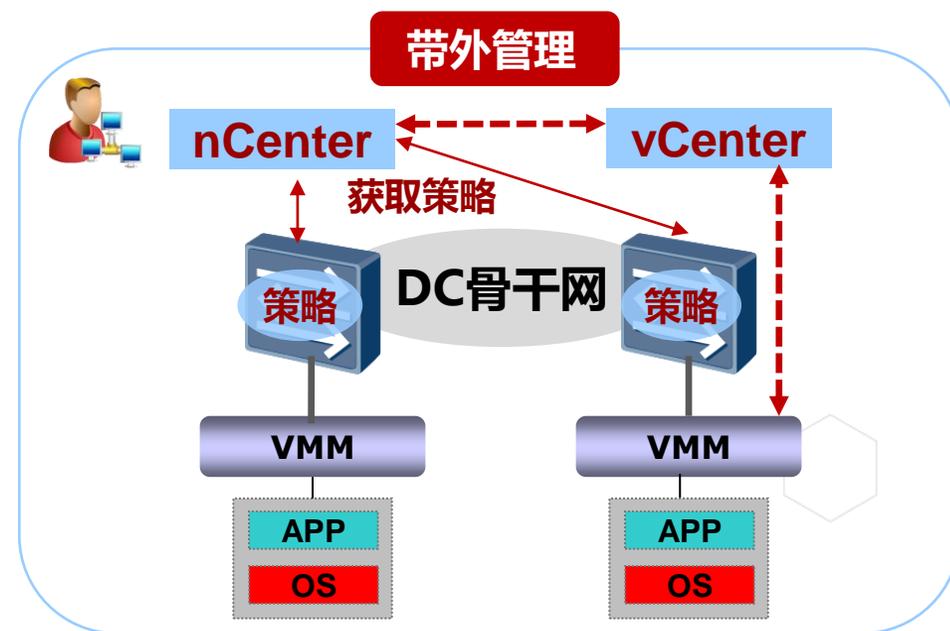
- 管理员在nCenter创建特定Port Profile的网络策略
- nCenter将Port Profile信息推送给vCenter
- vCenter将VM信息和网络策略设置到虚拟交换机
- nCenter从vCenter获取VM和虚拟交换机的关系，通过获取的拓扑信息，确定相连的接入交换机端口
- nCenter向接入交换机下发VM的信息，接入交换机从nCenter查询该VM的网络策略，完成接入交换机上的网络策略配置

- 由于虚拟机的发现是从vCenter获取的，这种方案称为带外管理方案
- 不同的虚拟机平台的API不同，nCenter需要适配各虚拟机平台的API
- 不需要虚拟机平台开发新的协议

# 虚拟感知解决方案选择：带内管理还是带外管理



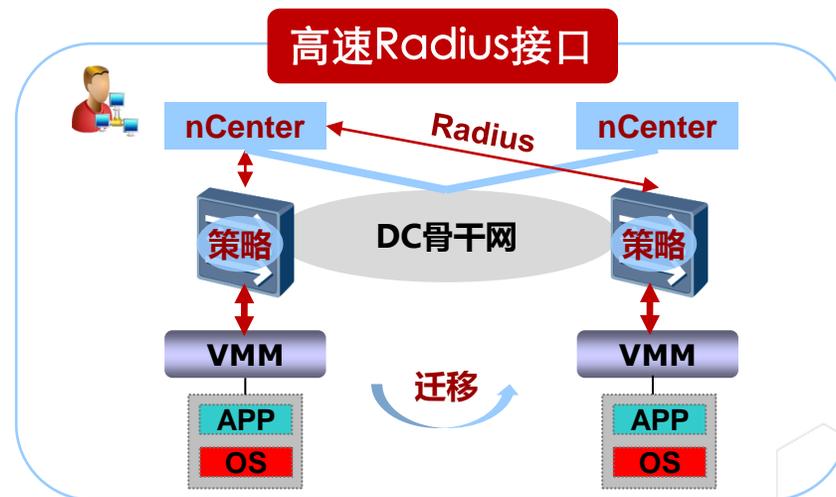
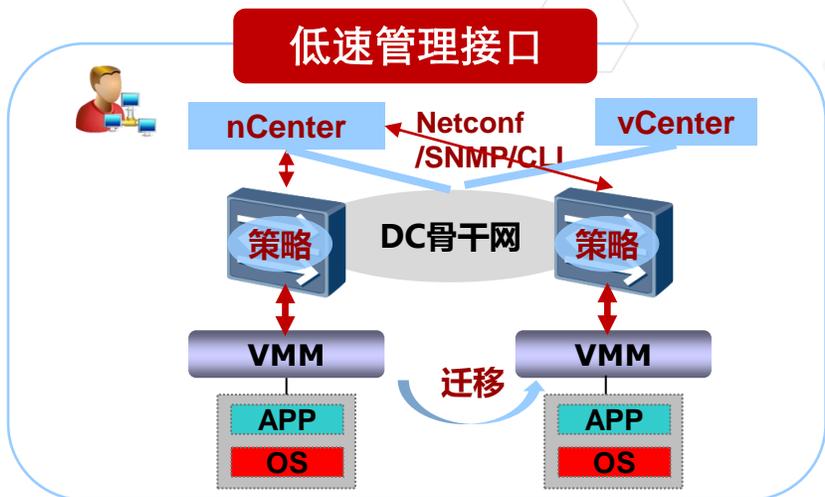
<b>缺点</b>	802.1Qbg目前还没有成为正式标准
<b>缺点</b>	虚拟机平台、交换机都尚未支持VDP协议
<b>缺点</b>	协议未具体规定和vCenter的接口



<b>优点</b>	兼容所有的虚拟机平台
<b>优点</b>	不需要虚拟机平台开发新功能，当前可获得
<b>缺点</b>	交换机从nCenter获取配置的接口没有标准

**带外管理协议方式，成熟度高，兼容性好，完全支持VM的动态管理！**

# 虚拟感知解决方案选择：交换机从nCenter获取策略的接口



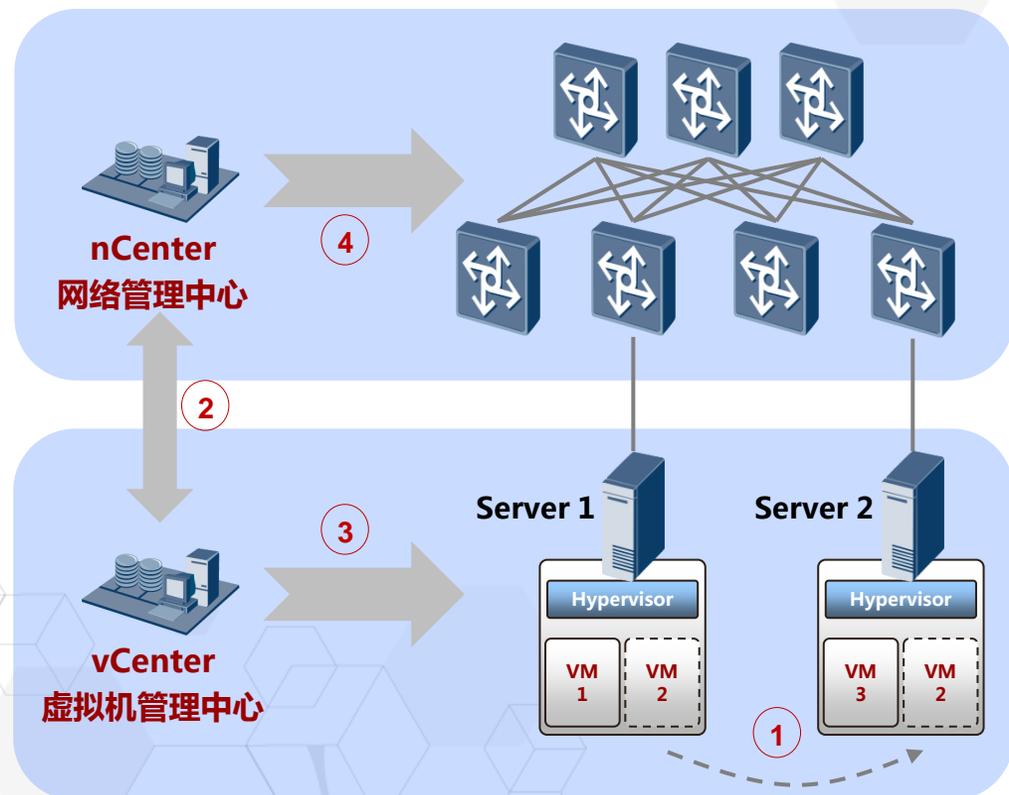
选择接口的决定因素：VM的上线速率



- 并发上线率 =  $(N/2) * 4 / 3 * 60$
  - N：物理服务器数量
  - 2：一半忙服务器向一半空闲服务器迁移
  - 4：每个物理支持4个VM迁移（测试数据，受限于带宽和CPU能力）
  - 3\*60：经验数据，每个VM的迁移时间为3分钟
  - ◆ 如果是1万台服务器，并发上线率 = 111个/秒
- Netconf: 每秒 < 20个VM上线      Radius: 每秒200个VM上线

**计算资源池要求网络支持VM的动态调度，Radius协议能够提供高速的VM策略下发！**

# 虚拟感知解决方案



## ① 虚拟机迁移

物理服务器1中的VM2虚拟机因业务调整需要迁移至物理服务器2中。

## ② vCenter和nCenter交互信息

vCenter将虚拟机迁移的相关信息告知nCenter。

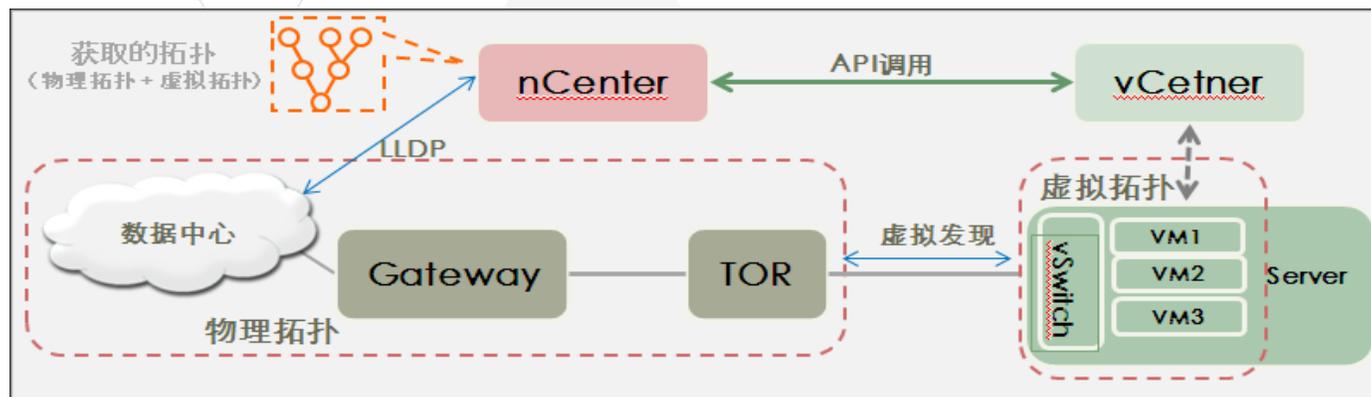
## ③ vCenter下发策略

vCenter生成迁移虚拟机原有的服务器策略，下发至相应的物理服务器。

## ④ nCenter下发策略

nCenter生成迁移虚拟机原有的网络策略，下发至相应的接入交换机。

# 虚拟感知解决方案：拓扑管理



## ① 物理TOPO：LLDP标准协议发现



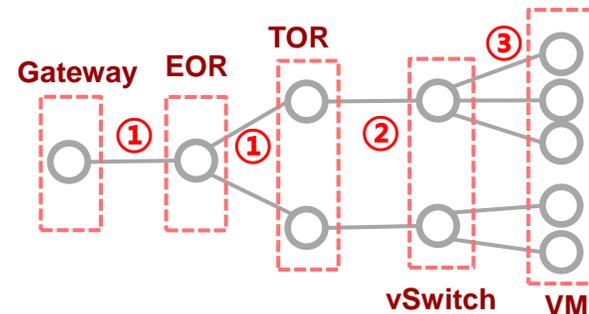
## ② TOR->vSwitch TOPO：LLDP/CDP/静态配置



## ③ vSwitch->VM：通过vCenter的API获取



## ④ nCenter管理整网TOPO：网络自动化部署的基础



# 虚拟感知解决方案：策略管理

**Basic Information**

\*

\* VLAN ID:

Create Profile

**Basic Information**

\* Profile name:

**QoS Information**

\* QoS name:  Policy direction:

\* CIR (kbit/s):  \* PIR (kbit/s):

**ACL Information**

\* ACL name:

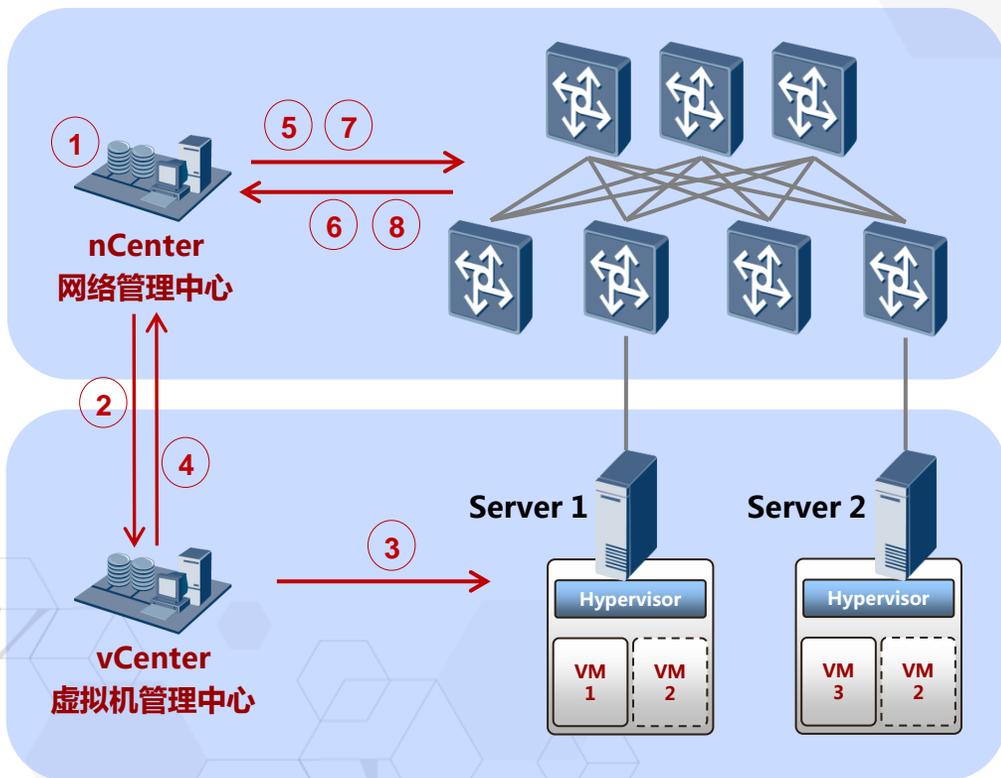
## 虚拟机（VM）策略概念

- 包括静态策略和动态策略。
- 静态策略包括VLAN、ACL、QoS等。
- 动态策略包括DHCP Snooping、MFF等。

## 策略管理

- 在VM启动前网络管理员在nCenter配置好策略模板。
- nCenter将策略推送给vCenter，显示为PortGroup。
- 服务器管理员通过vCenter创建VM时把VM和策略模板进行绑定，并把绑定关系通知给nCenter，使nCenter能够跟踪和维护VM与策略模板之间的关系。

# 虚拟感知解决方案：虚拟感知（创建VM）

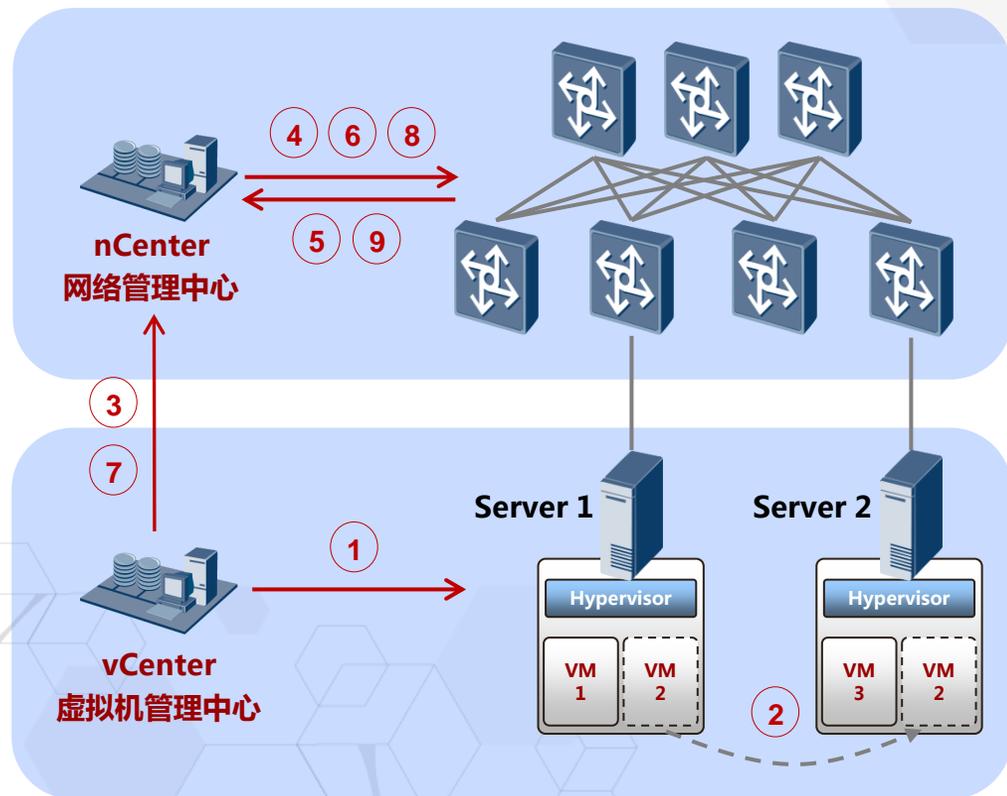


- ① nCenter上创建VM的网络属性模版。
- ② nCenter将网络属性模版推送给vCenter。
- ③ vCenter创建并启动VM，并将服务器内网络属性配置到vSwitch。
- ④ nCenter通过vCenter感知VM上线，保存VM和策略模板的绑定关系。
- ⑤ nCenter通过netconf action机制通知TOR交换机VM上线，上线信息中包含VM的MAC地址、VLAN信息、应用的策略模板ID。
- ⑥ TOR通过Radius协议向nCenter申请VM策略（ACL、QoS）。
- ⑦ nCenter中的Radius Server响应申请，将VM绑定策略应答给TOR，TOR接收到Radius的消息，解析出VM的策略，完成转发配置。
- ⑧ VM如果通过DHCP申请IP，则TOR通过DHCP Snooping获取到VM的IP后，通过Radius协议上送给Radius Server（nCenter）保存。

\* VM如果是静态IP，建议修改为通过DHCP申请IP，同时在DHCP Server上为这个VM的MAC设置固定IP。

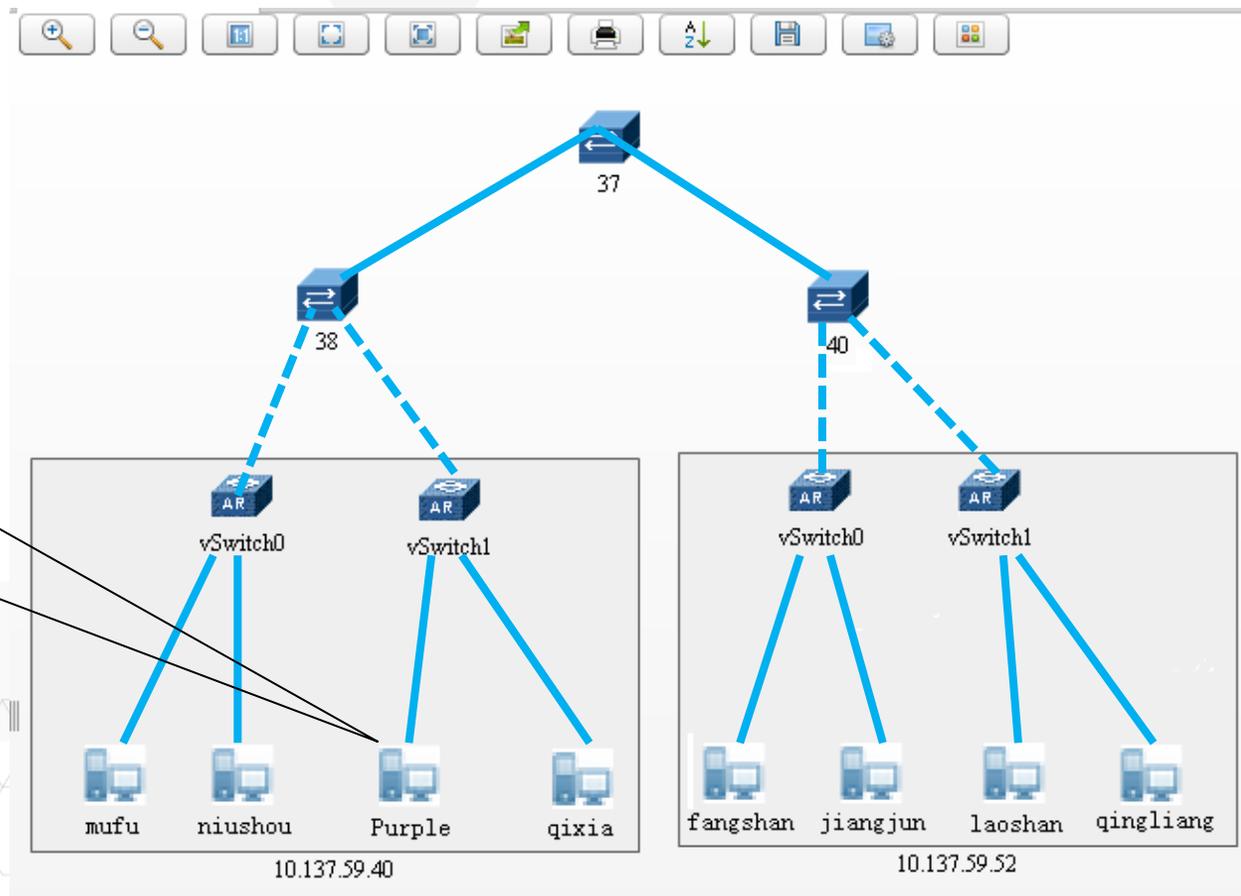
这种方式的优点在于能够通过DHCP Snooping实现安全策略的动态迁移。

# 虚拟感知解决方案：虚拟机迁移过程



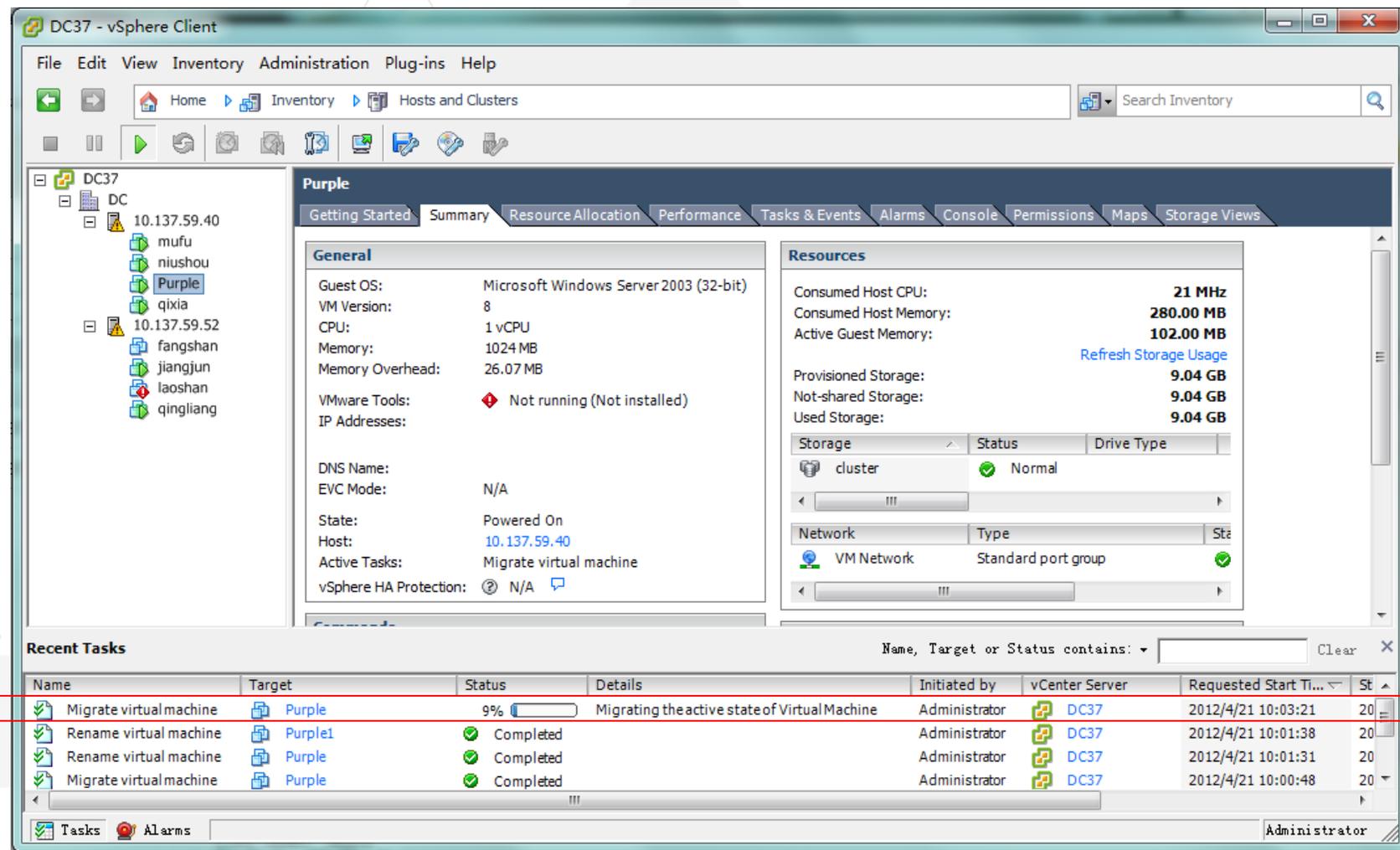
- ① vCenter启动VM迁移。
- ② 进行VM迁移。
- ③ vCenter通过API告知nCenter迁移开始。
- ④ nCenter通知目的TOR交换机VM上线，上线信息中包含VM的身份信息VM的MAC地址、VLAN信息、应用的策略模板ID。
- ⑤ 目的TOR通过Radius协议向nCenter申请VM策略（ACL、QoS、DHCP绑定表）。
- ⑥ nCenter中的Radius Server响应目的TOR的申请，将VM绑定策略应答给目的TOR，目的TOR收到后，解析出VM的策略，完成转发配置。
- ⑦ vCenter通过API告知nCenter迁移完成。
- ⑧ nCenter通知原TOR交换机VM下线。
- ⑨ 原TOR接收到下线通知，删除本地策略的同时，通过Radius的用户下线接口通知Radius服务器更新用户在线状态。

# 虚拟机迁移过程：迁移前拓扑



准备迁移的虚拟机  
**Purple**  
目前在接入交换机38  
连接的物理服务器  
10.137.59.40上

# 虚拟机迁移过程：在vCenter上发起迁移



将虚拟机  
**Purple**  
从10.137.59.40  
迁移到10.137.59.52  
当前状态：刚刚启动

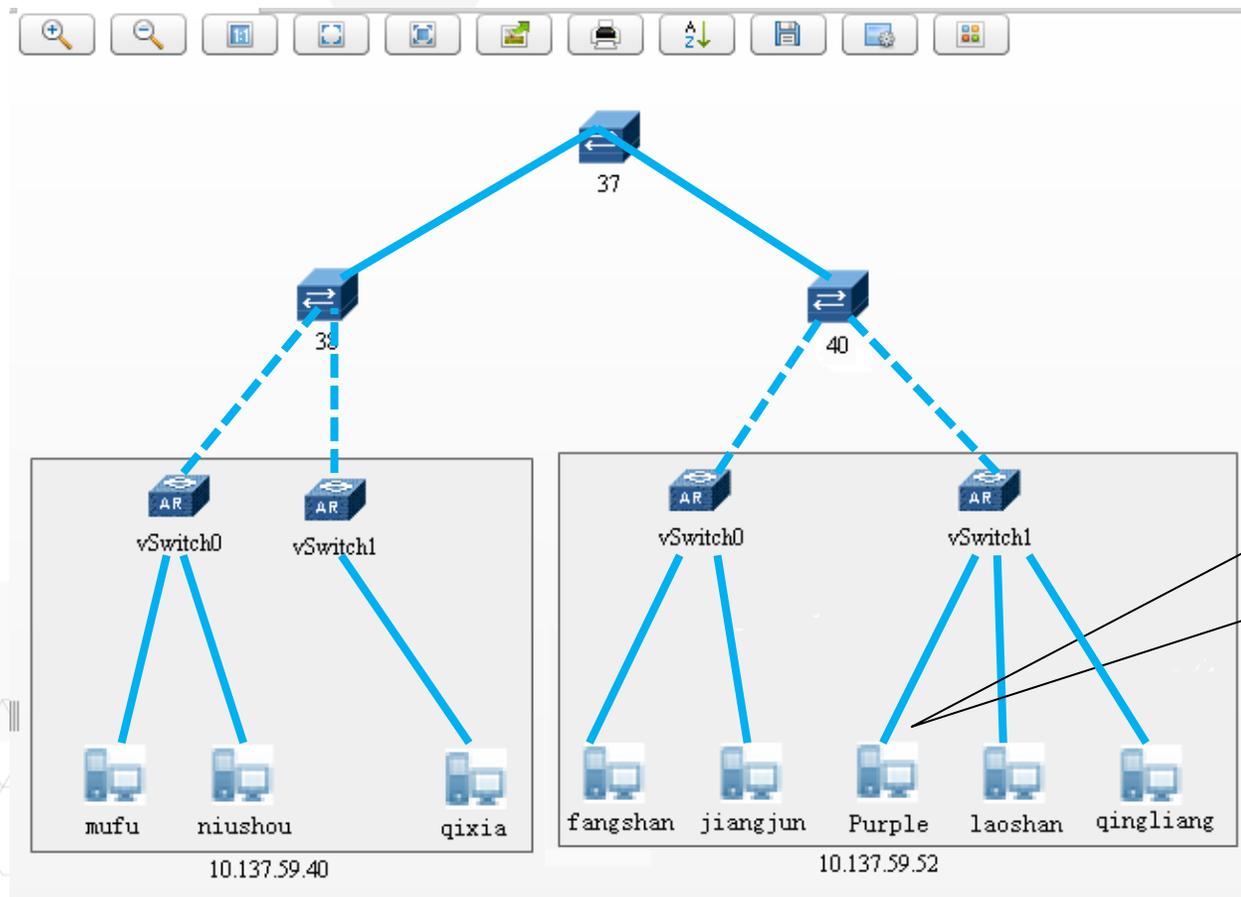
# 虚拟机迁移过程：在vCenter查看迁移进度

The screenshot shows the vSphere Client interface for a host named 'DC37'. The left sidebar shows a tree view of hosts and clusters, with 'Purple' selected under the IP address 10.137.59.52. The main pane displays the 'Summary' tab for the VM 'Purple', showing details like Guest OS (Microsoft Windows Server 2003), VM Version (8), CPU (1 vCPU), Memory (1024 MB), and State (Powered On). The 'Resources' section shows CPU and memory usage. At the bottom, the 'Recent Tasks' table lists several migration tasks, with the first one highlighted by a red box.

Name	Target	Status	Details	Initiated by	vCenter Server	Requested Start Time	St
Migrate virtual machine	Purple	Completed		Administrator	DC37	2012/4/21 10:03:21	20
Rename virtual machine	Purple1	Completed		Administrator	DC37	2012/4/21 10:01:38	20
Rename virtual machine	Purple	Completed		Administrator	DC37	2012/4/21 10:01:31	20
Migrate virtual machine	Purple	Completed		Administrator	DC37	2012/4/21 10:00:48	20

将虚拟机  
**Purple**  
从10.137.59.40  
迁移到10.137.59.52  
当前状态：迁移完成

# 虚拟机迁移过程：迁移后拓扑



虚拟机  
**Purple**  
已经迁移到接入交换机40  
连接的物理服务器  
10.137.59.52上

# 虚拟感知解决方案规格



支持虚拟化管理平台VMware

VLAN、ACL、QoS等静态配置自动部署

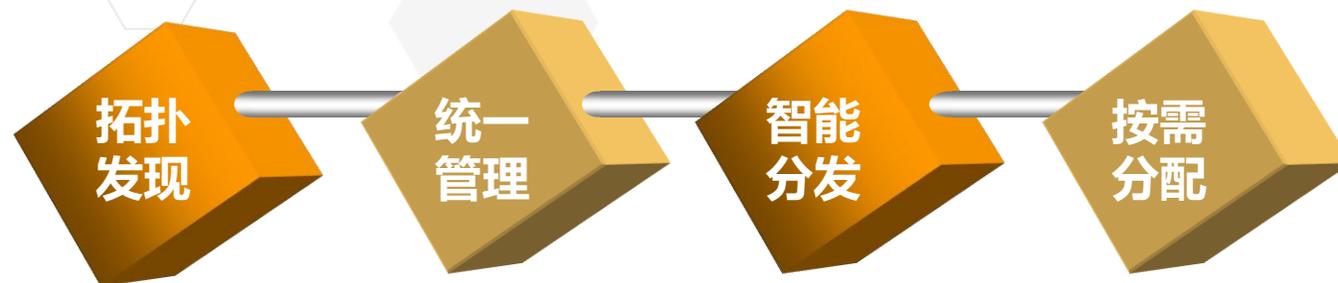
nCenter内置高性能Radius服务器，每秒200个虚拟机上线

支持包括vSwitch在内的网络拓扑收集、显示

DHCP Snooping、MFF等动态配置自动部署

管理虚拟机、虚拟交换机、物理交换机网元2万个

# 虚拟感知解决方案亮点



**拓扑信息**  
自动收集；  
拓扑展示  
自动更新。

**物理网络**  
虚拟网络；  
全网可见  
统一管理。

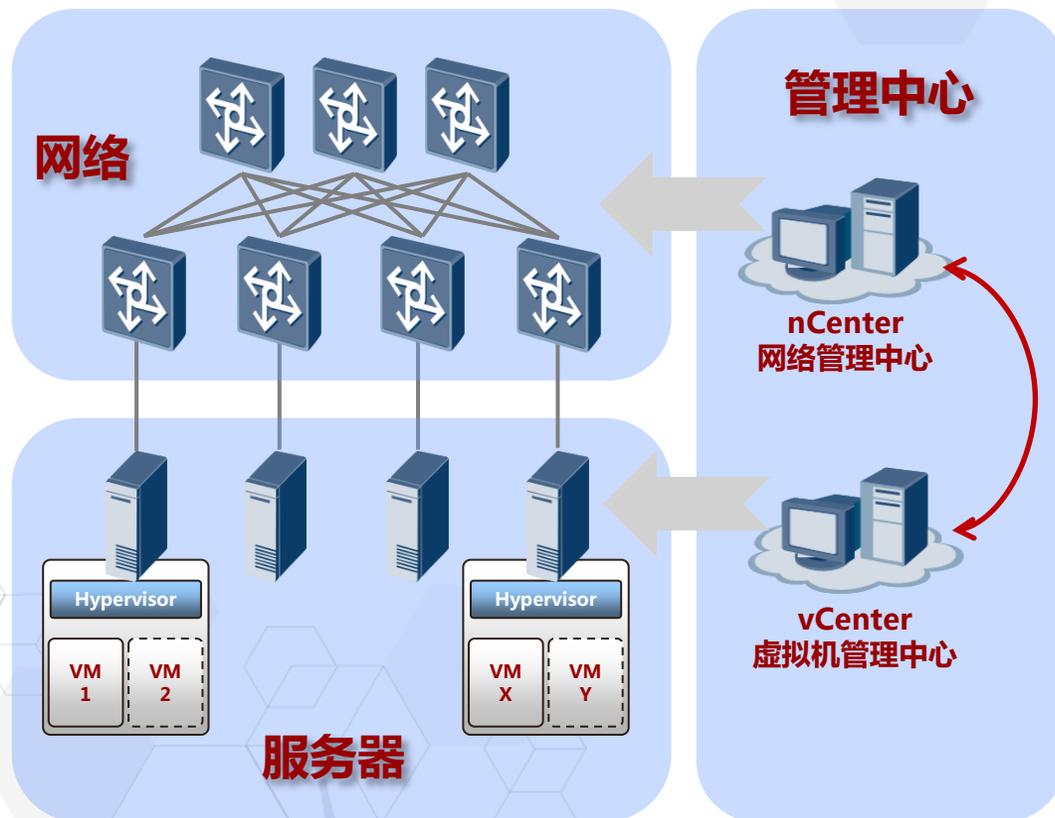
**网络策略**  
自动生成；  
网络配置  
自动下发。

**网络资源**  
实时监控；  
网络资源  
及时回收。

# 目录

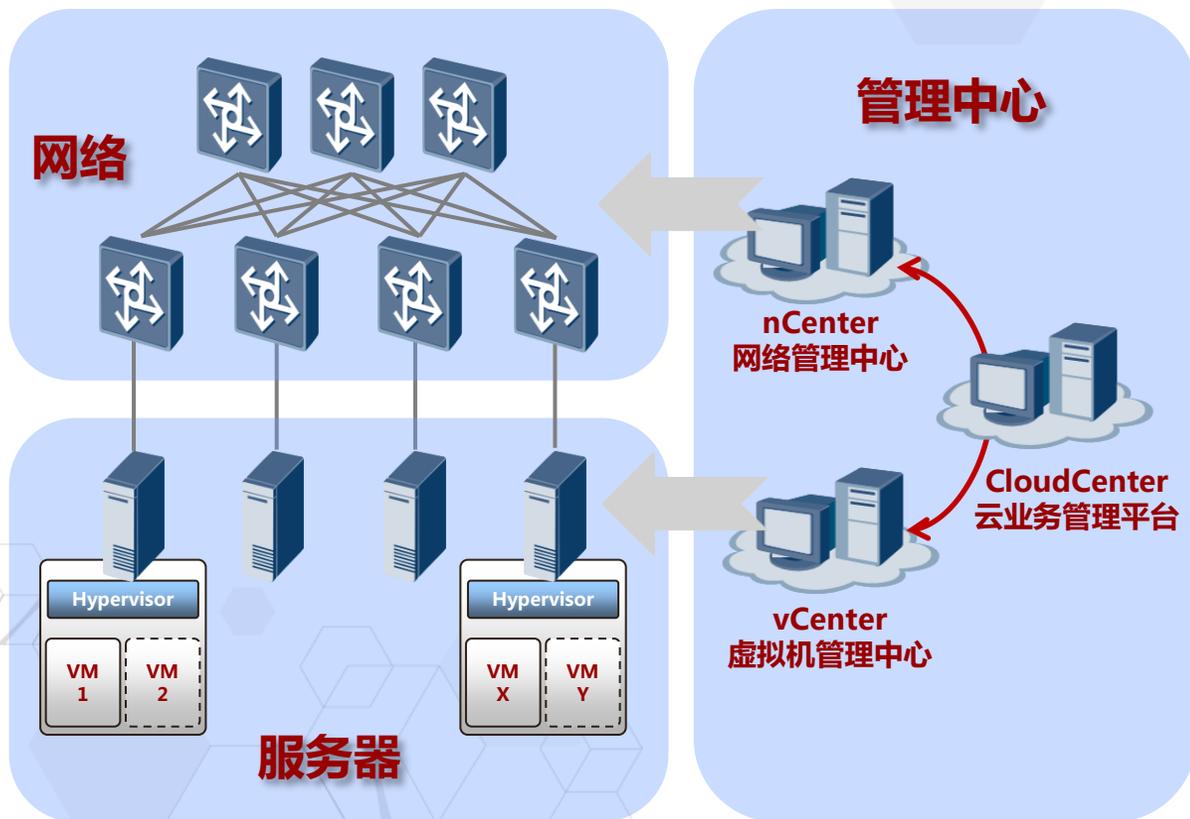
- 服务器虚拟化对数据中心网络的挑战
- 华为虚拟感知解决方案技术原理
- 华为虚拟感知解决方案应用部署

# 虚拟感知应用部署



- vCenter是虚拟机管理中心。
- nCenter是网络管理中心。
- vCenter和nCenter部署在数据中心的管理中心。
- 在网络中交换机上进行基本的网络配置：建立VLAN或TRILL转发网络；使能LLDP等协议；建立管理通道。
- 在TOR交换机上配置使能虚拟感知。
- 在nCenter上配置VM的网络模版。
- 配置nCenter和vCenter的配合关系。
- 用户在vCenter上进行VM的创建、迁移的管理。

# 虚拟感知应用部署：通过云业务管理中心集中管理



- 增加CloudCenter云业务管理中心。
- 在网络中交换机上进行基本的网络配置：建立VLAN或TRILL转发网络；使能LLDP等协议；建立管理通道。
- 在TOR交换机上配置使能虚拟感知。
- 用户在nCenter上配置VM的网络模版。
- 配置nCenter和vCenter的配合关系。
- 用户在CloudCenter上进行VM的申请、扩容等管理，CloudCenter调用vCenter、nCenter完成部署。



**HUAWEI**

**Huawei Enterprise *A Better Way***