思科超融合HyperFlex,一个跨界竞争的时代

张亦安

演讲内容

- 超融合及其应用场景
- HyperFlex 市场前景及成功案例
- HyperFlex 产品简介、技术详解和优势分析
- 业界动态和竞争简析

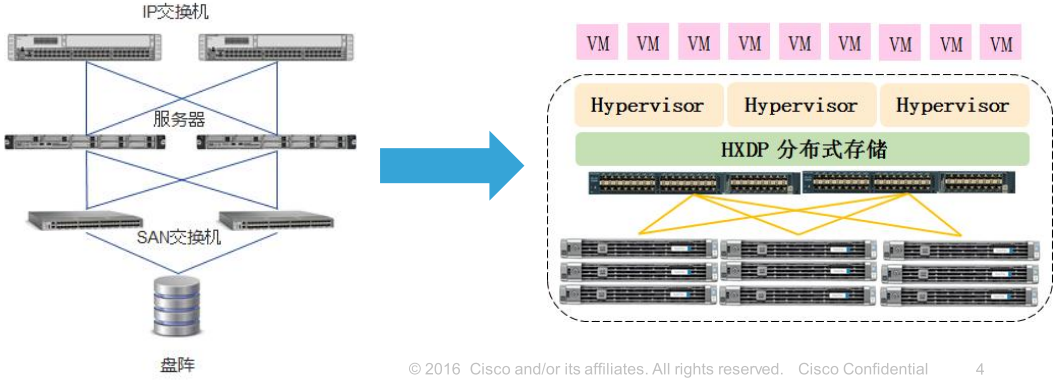
© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

超融合及其应用场景

什么是超融合？它的市场驱动力是什么？

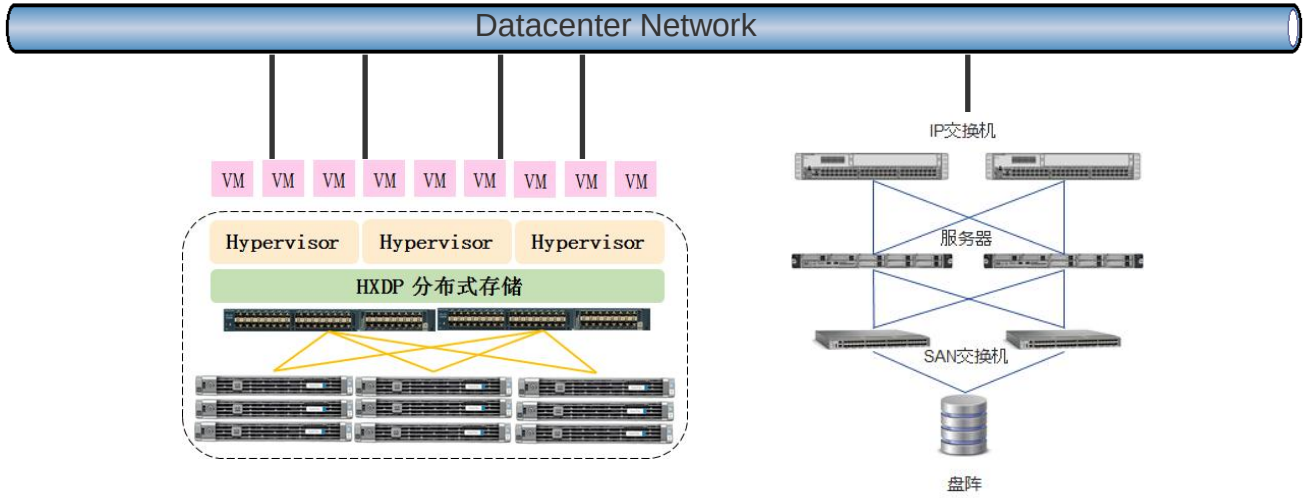
- 什么是超融合？
  - 共享计算、存储、网络资源的平台
  - 软件定义存储、软件定义计算、软件定义网络

- 超融合市场驱动力：方便
  - 运维简单，管理方便
  - 部署简单，变更方便
  - 方案简单，调整方便



© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

未来的IT数据中心架构：简化、灵活、软件定义



- 超融合HyperFlex
  - 敏捷业务、快速简单；如测试开发、教学办公
  - 云化业务，高效低成本；如服务器虚拟化、科研云
  - 弹性业务，按需扩展；如分支机构、VDI
- 传统架构
  - 固定任务，功能标准；如Core banking
  - 大型任务，应用单一；如HPC、游戏
  - 应用限制，尚需时间；如证券交易

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

HyperFlex 典型应用场景

虚拟桌面（VDI）

服务器虚拟化

测试开发

分支机构

云计算

- 弹性需求：如服务器虚拟化、测试开发、云计算
- 简化管理需求：如分支机构、工厂、商场、酒店、4S店
- 横向扩展需求：如VDI、APP接入

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

HyperFlex领先地位及成功案例

思科在集成系统(融合及超融合)市场处于领先地位

sco and/or its affiliates. All rights reserved. Cisco Confidential

HyperFlex 全球成功案例，500+

金融服务

政府

教育

Software / Services

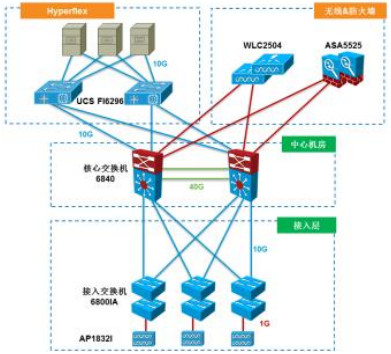
医疗保健

其它

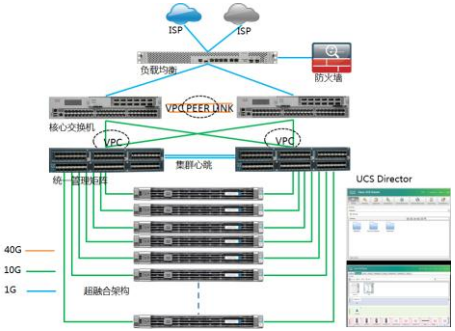
© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

HyperFlex在大中国区成功案例

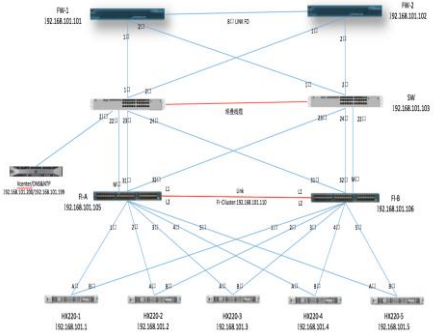
某纺织集团第2代IT系统



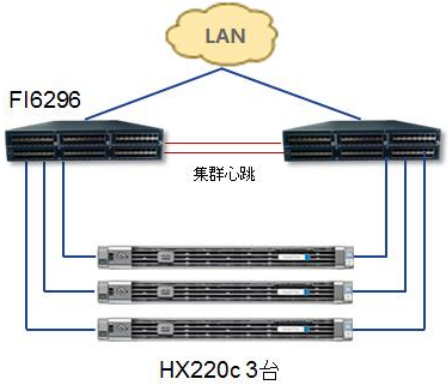
某上市医药公司云计算平台



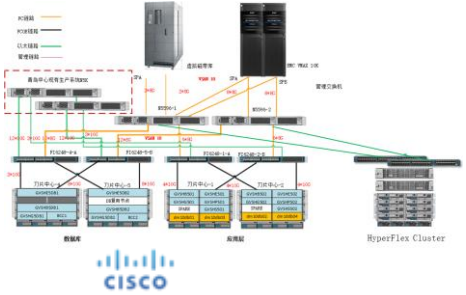
某大宗商品交易平台



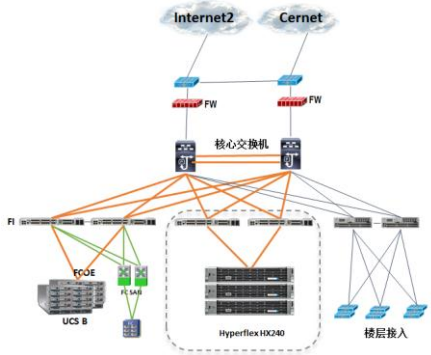
某大型金融公司敏捷业务平台



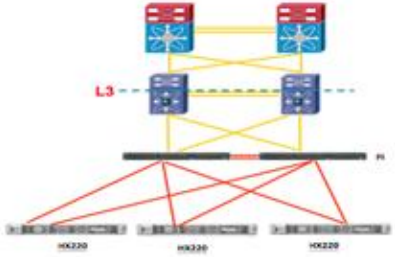
某家电集团测试开发平台



某大学教学办公IT支撑平台

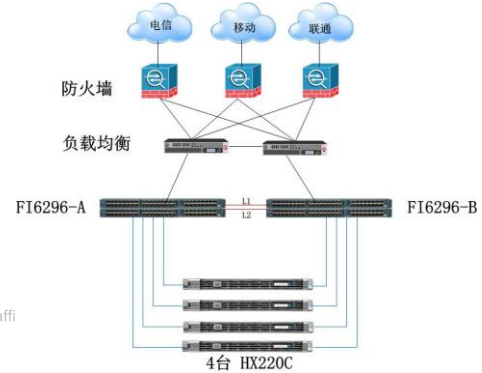


某大学虚拟桌面



© 2016 Cisco and/or its affi

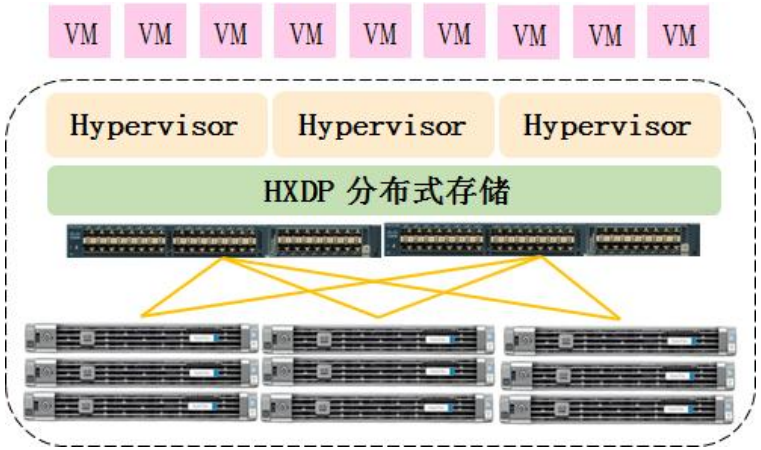
某证券公司网站虚拟化



HyperFlex 产品简介

HyperFlex 产品简介

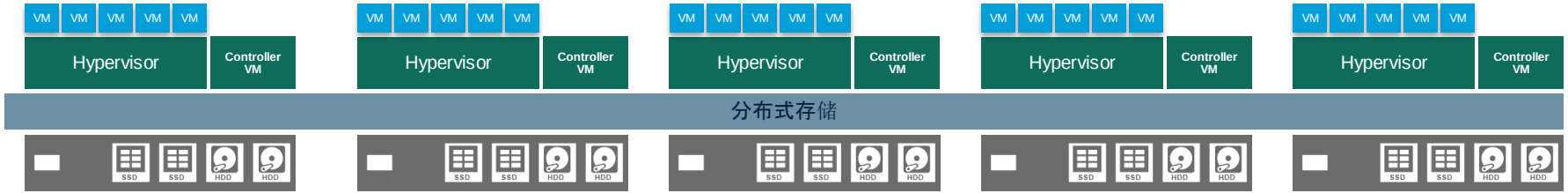
UCS + FCoE交换机 + 分布式存储软件HXDP + 虚拟化软件 = HyperFlex



- HyperFlex 硬件是标准UCS
  - 2台光纤交换互联矩阵(FI)
  - 3 台~ 128台UCS服务器
  - 20TB ~ 1760TB存储容量
- HyperFlex 软件是商用软件
  - HXDP 分布式存储软件
  - 服务器虚拟化软件(Hypervisor)

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

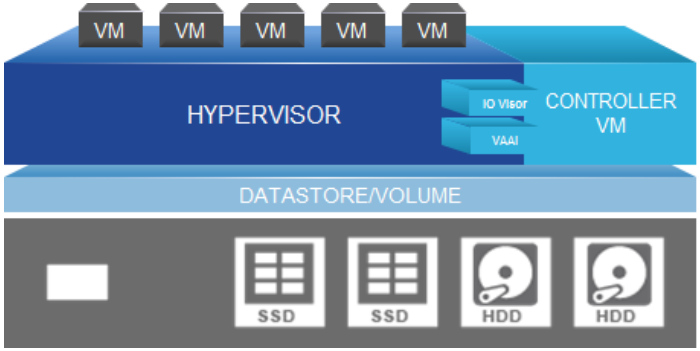
HXDP分布式存储软件基本原理



- Hypervisor和Controller VM的软件都安装在HyperFlex的64G SD卡上，每台物理服务器都要安装
- Hypervisor启动pass through直通模式，所有物理盘包括SSD和磁盘的控制都是由Controller VM进行控制。所有Controller VM组成Cluster，逻辑上整个集群是一块本地硬盘

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

HXDP分布式存储软件基本原理(续)



- Hypervisor里安装HXDP驱动和IO Visor和VAAI
  - IO Visor 负责把整个分布式存储像本地硬盘一样呈现给Hypervisor
  - VAAI 负责传递配置和管理命令
- Hypervisor再从这块逻辑上像本地盘一样划分空间创建Datastore，然后就可以创建虚机VM了
- 虚机VM的IO先传给Hypervisor，再传给IO Visor，再给到Controller VM，再落到物理硬盘上

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

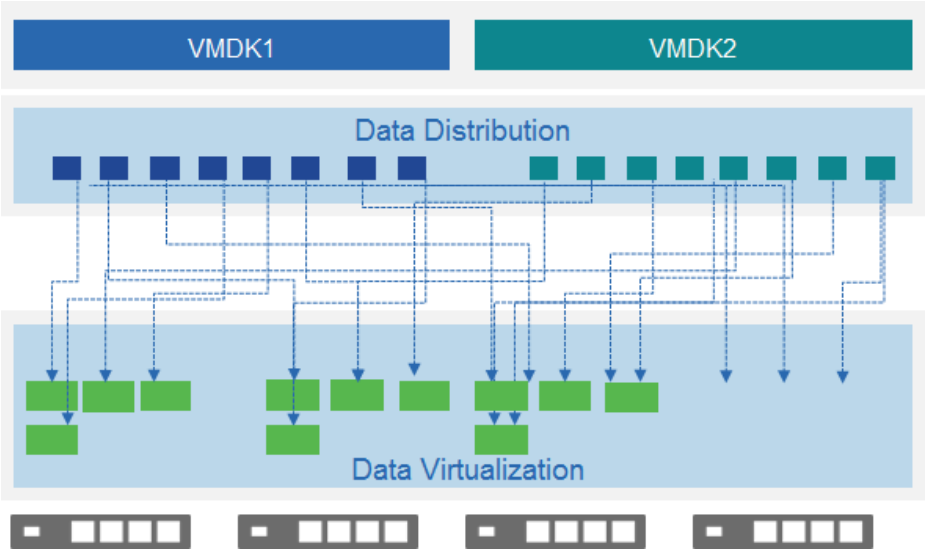
HXDP分布式存储中SD卡、SSD盘、磁盘的用途

存储介质	数量	用途	描述
64G SD卡	2块	存放Hypervisor软件 存放Controller VM虚机	镜像，防止介质损坏
120G SSD盘	1块	IO的日志 数据的指针	2/3拷贝，镜像到其它服务器节点上
480GB/1.6TB SSD盘	1块	写数据的缓存 读数据的缓存	写数据的缓存2/3拷贝，镜像到其它服务器节点上 读数据的缓存不需要镜像
1.2TB 一万转SAS磁盘	6-23块	存放数据	2/3拷贝，镜像到其它服务器节点上

- 数据(Data)：存放在存储介质上的0/1
- 元数据(Metadata)：也叫指针，描述数据在存储介质上的物理位置
- 日志(Log)：描述写操作的状态，主要是因为缓存技术出现和使用，应用于非正常关机(如掉电)后的一致性恢复

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

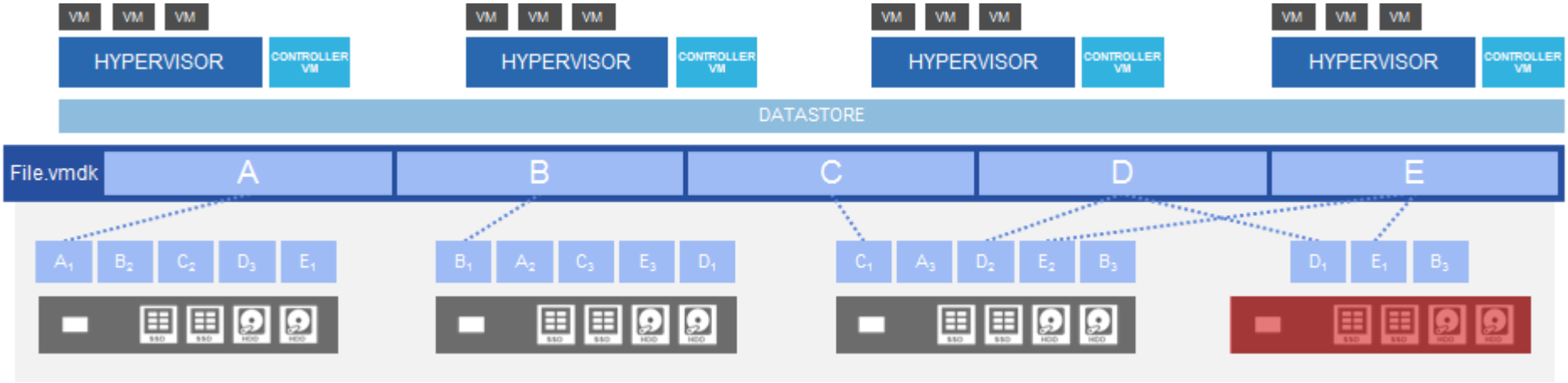
HXDP数据虚拟化:全局虚拟化



- 条带化不局限于本地磁盘，跨越整个服务器
- 虚机VM能使用到整个集群的容量

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

HXDP中的数据保护

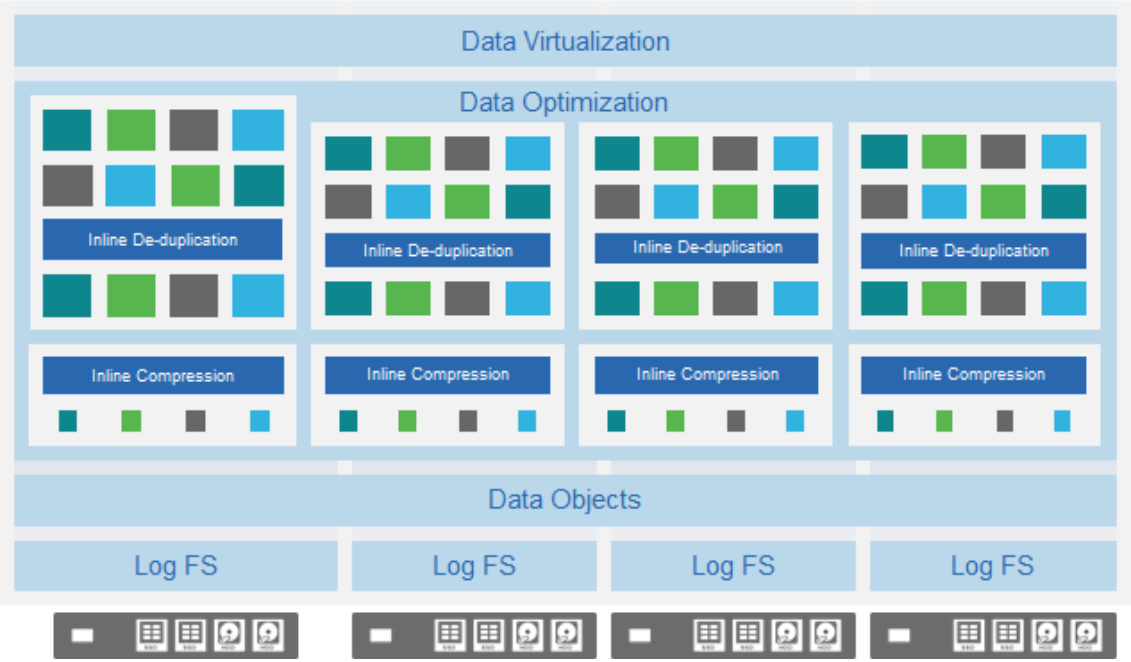


- 所有写数据都是2/3份镜像拷贝，分布在不同服务器节点上
- 64GB SD卡损坏，自动访问镜像SD卡，无任何影响
- SSD盘损坏，自动访问镜像SSD；应用不中断，IO延迟加大性能下降
- 1.2TB磁盘损坏，自动访问镜像磁盘；应用不中断，IO性能不下降
- 服务器宕机，虚机VM自动在其它服务器上重启，应用要中断约5-15分钟，IO性能不下降

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

HXDP持续数据优化

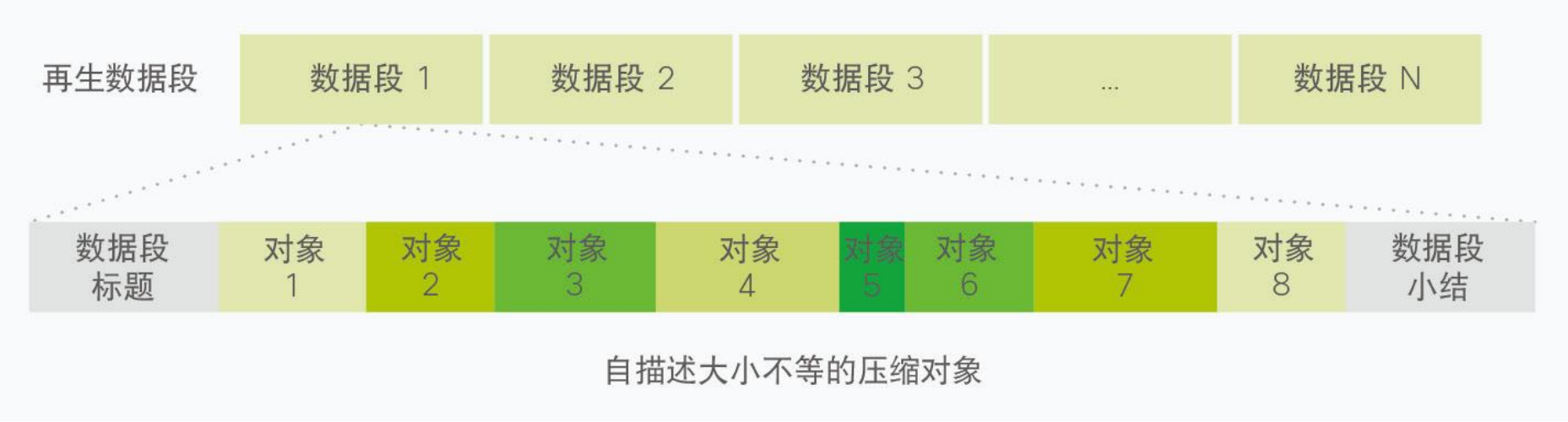
- 结构化日志型(Log-structured)文件系统让数据优化更有效
  - 重复数据删除和压缩都是在线的，只有1次写，减少50% IO消耗，延长SSD寿命一倍



All rights reserved. Cisco Confidential

结构化日志，对象存储

- 结构化日志技术来自Google，对象存储最新存储技术，其优点：
  - 支持海量扩展，搜索速度快，指针性能衰减小
  - 顺序布局，提高闪存寿命，提高磁盘HDD的性能
  - 对压缩和快照操作或整体性能的性能影响很小或没有影响
  - 故障恢复速度快



分布式存储的演进

Google File System
(2003)

Ceph
(2006)

HDFS
(2007)

Swift
(2010)

Nutanix
(2009)

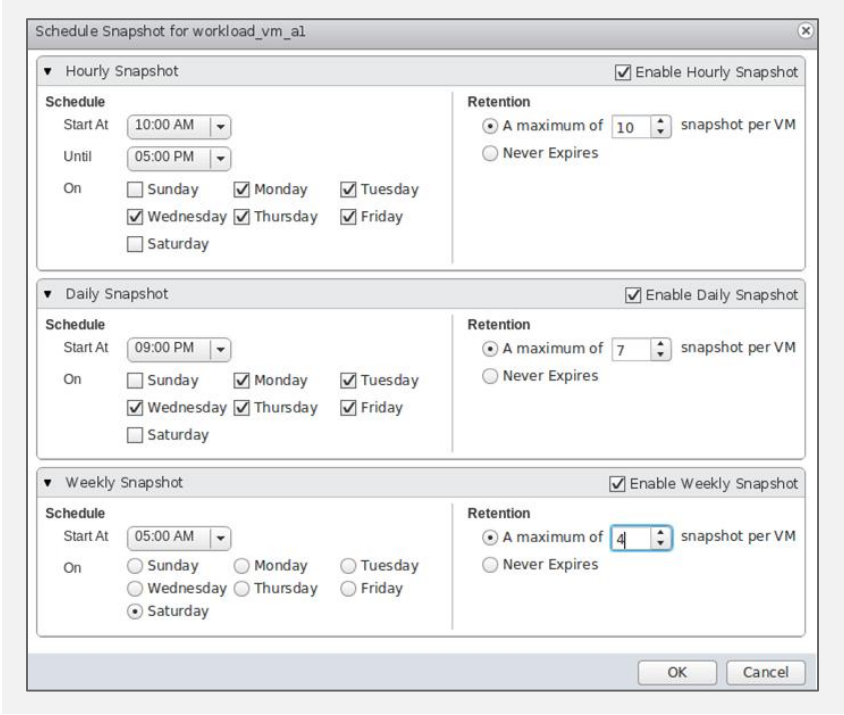
HXDP
(2012)

VSAN
(2014)

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

HXDP快照，快速灵活

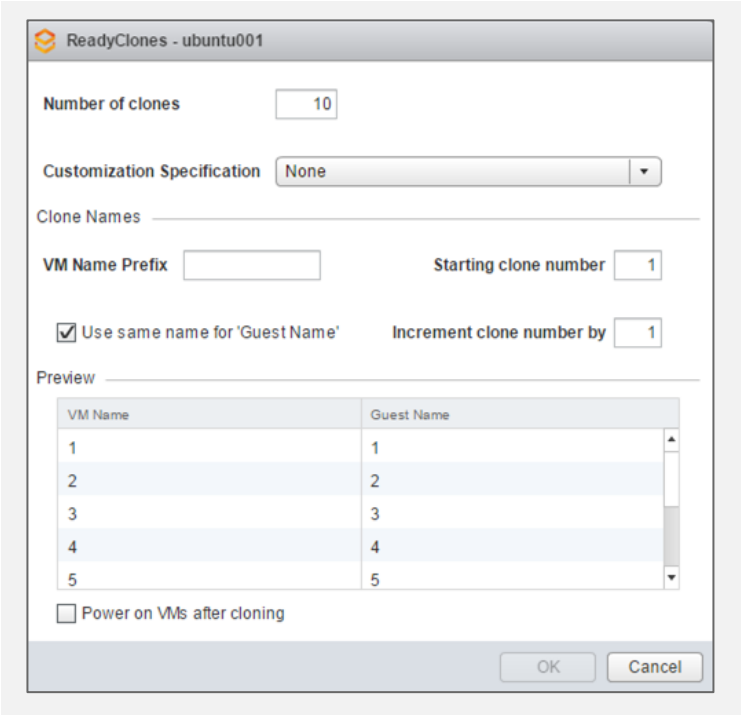
- 基于指针的快照
  - 节省空间
  - 快速生成和删除
- 快照颗粒度
  - VM层面或VM文件夹层面均可
- VAAI集成
  - 静默以保证一致性
- 管理上用 vCenter Snapshot Manager
- 支持自动策略
  - 可以排计划，可以设定保留时间



© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

HXDP 快速虚机克隆

- 基于指针的可写快照 (实时克隆)
- VAAI 集成
- VM 层面
- 最多256个克隆



© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

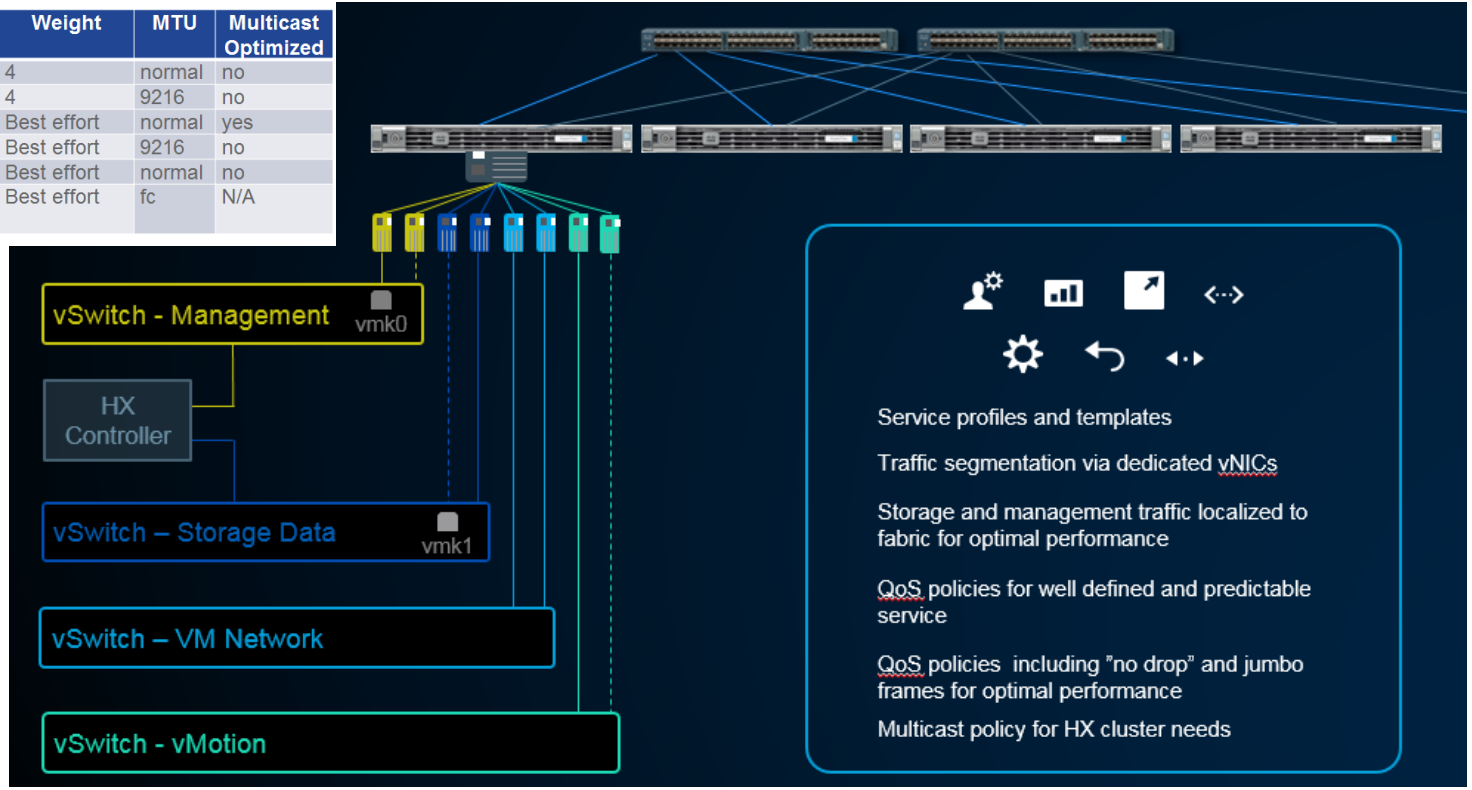
HyperFlex网络技术:高带宽、低延迟

- 低延迟: 端到端延迟为 3.2 微秒, 比普通交换机快50%
  - 低网络延迟对分布式存储的性能的作用很大
- 高带宽:最大带宽单台960 Gbps, 两台1920 Gbps
  - 高带宽在数据重构时候的作用很大

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

HyperFlex网络技术:QoS保障分布式存储的稳定

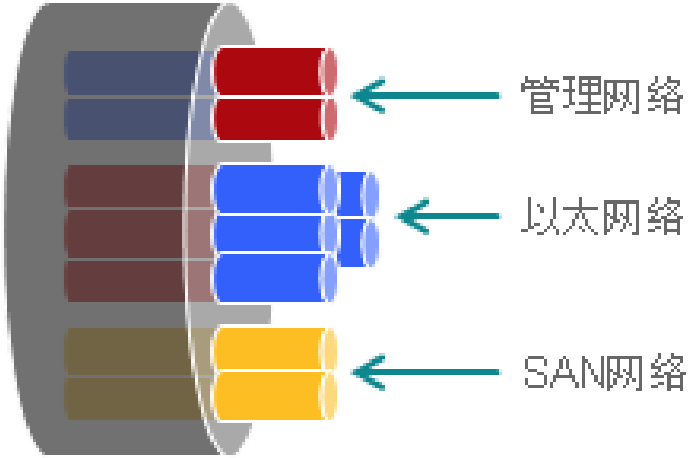
Priority	Enabled	CoS	Packet Drop	Weight	MTU	Multicast Optimized
Platinum	✓	5	Yes	4	normal	no
Gold	✓	4	Yes	4	9216	no
Silver	✓	2	Yes	Best effort	normal	yes
Bronze	✓	1	Yes	Best effort	9216	no
Best Effort	✓	Any	Yes	Best effort	normal	no
Fibre Channel	✓	3	No	Best effort	fc	N/A



© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

HyperFlex网络技术:FCoE

- 同时支持IP和SAN网络，从而实现对传统SAN存储连接
- 一次布线，灵活配置，简化管理
- 减少50%布线，



© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

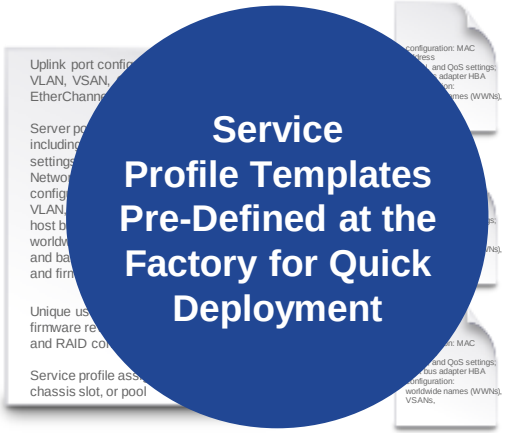
HyperFlex服务器技术:UCS无状态计算Service Profile

Server SME

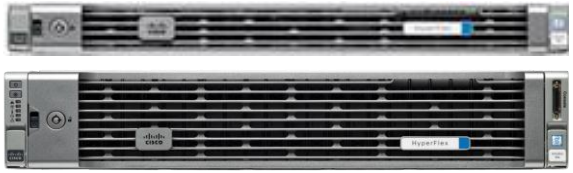
Network SME

Storage SME

- Server Policy
- Storage Policy
- Network Policy
- Virtualization Policy
- Application Profiles



工厂生产就
配置好



1 各方面专家一起确定好参数

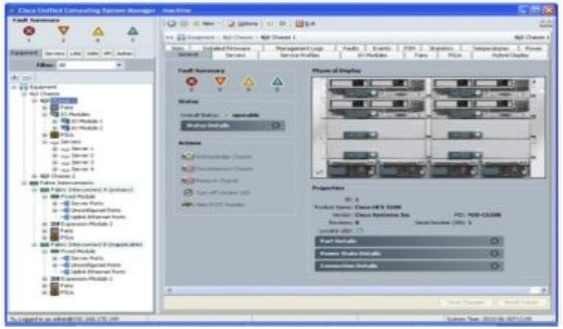
2 参数用来创建Service Profile模板

3 模板用来生成具体Service Profile

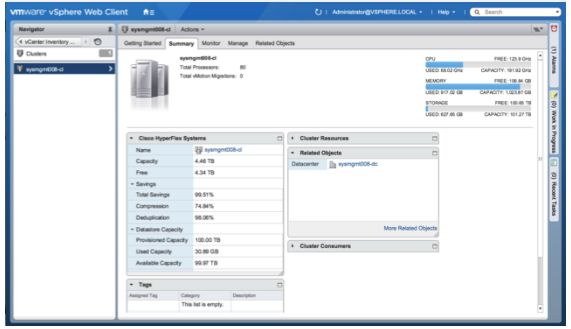
4 UCS服务器自动关联Service Profile和硬件配置

HyperFlex 管理技术:集成管理和嵌入式管理

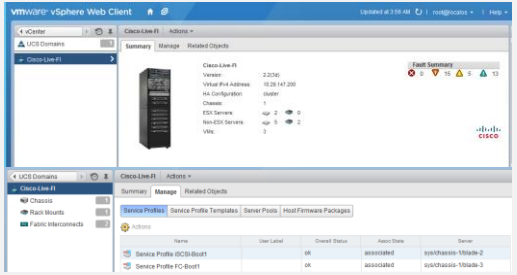
- UCS Manager 管理软件已经部署到全球 50,000 个用户
- 服务器、网络配置通过预定义的 Service Profiles进行部署



- 实施虚拟的创建、配置、克隆和快照通过 vCenter
- 没有额外的管理入口，没有额外的学习曲线



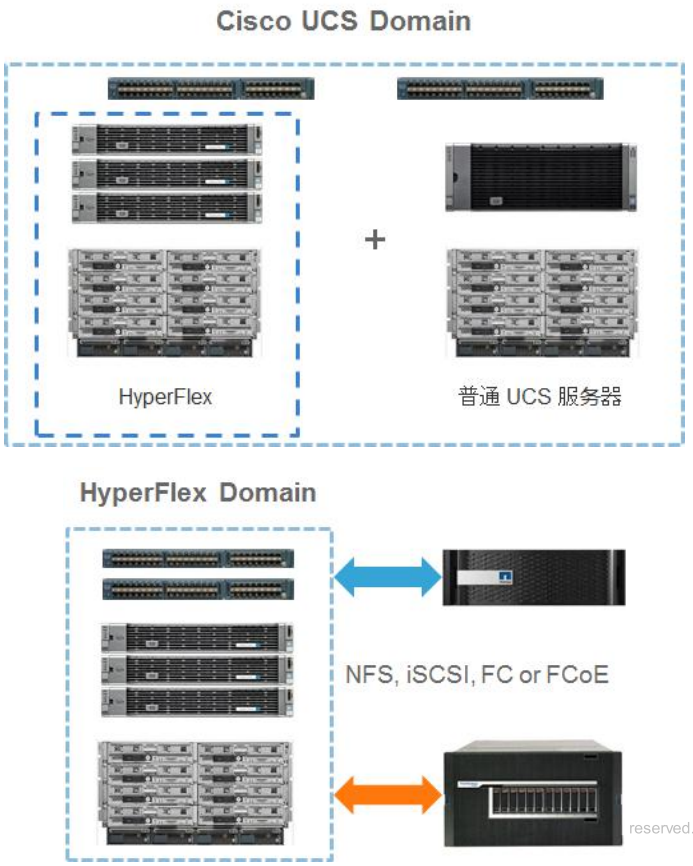
- UCS Manager vCenter Plug-in
- 管理 UCS Service Profiles
- 升级 Server Firmware
- KVM launch of console



© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

HyperFlex扩展性：可以单独扩，可以一起扩

- 计算容量都扩
  - 加HyperFlex节点即可
- 只扩容量
  - 加盘即可
  - 增加外置SAN存储
- 只扩计算
  - 加普通UCS服务器即可



HyperFlex 技术优势

1. **完整全面的商用系统，自主研发一站服务**
   - 交换机、服务器、存储都配全了，自主研发统一服务

2. **集成的低延迟高带宽的网络，端到端的QoS保障**
   - FCoE 同时提供IP和FC SAN网络能力
   - 3.2微秒低延迟，1920Gb的带宽，端到端QoS保障

3. **第二代存储技术，持续的数据优化**
   - 在线Dedup和压缩，数据只写1次，节省50%带宽消耗，闪盘寿命延长1倍
   - 全局容量，自动平衡和优化

4. **独立扩展，可单独扩盘也可单独扩计算能力**
   - 可以连接传统的UCS服务器，机架刀片均可
   - 可以连接传统的SAN存储，IP SAN、NAS、FC存储均可

5. **部署快捷，管理简单，Cloud Ready**
   - HyperFlex部署快捷，管理简单，并可以平滑升级成云计算

cisco

业界动态和竞争分析

HyperFlex是第二代超融合产品

第一代超融合
只融合了服务器和存储

NUTANIX
COMPUTE. STORAGE. COMPLETE.

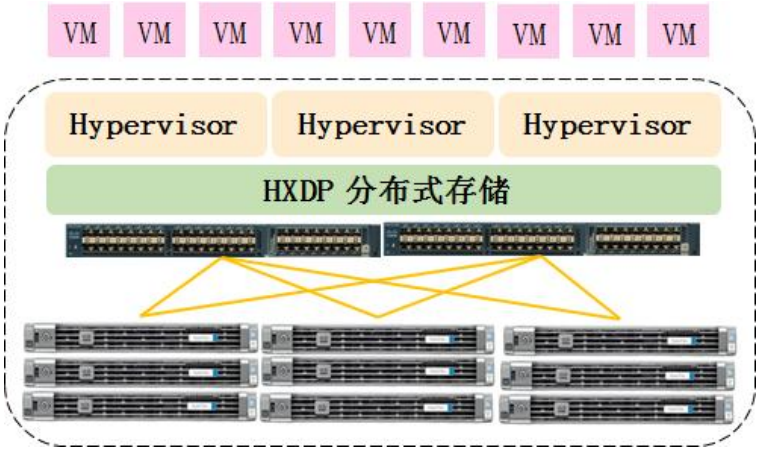
simplivity™

vmware®
Virtual SAN

VXRAIL

cisco

第二代超融合HyperFlex
融合了服务器、存储、网络



© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

Nutanix很危险，一直处于亏损状态的公司

- 宏观角度
  - Nutanix 2016年亏损1.68亿美金，一直在推迟上市，现在这种经济情况不乐观
  - Cisco有了HyperFlex，Dell买了EMC，HPE自顾不暇，联想情况也很不妙，Nutanix卖身的前途也不光明
  - Nutanix服务器OEM成本不低，在美国上市利润率很重要，Nutanix价格无法便宜
  - 销售和服务人数、资源和能力严重短缺
  - 联想OEM Nutanix同时，又有一个很弱的基于OpenStack开源的超融合产品AIO
- 产品和技术角度
  - 分布式存储技术中没有原生态的Dedup和压缩
  - 服务器、网络、SAN方面没有任何技术能力，网络和SAN方面连服务能力都没有
  - 虚拟化由KVM改的，恰恰说明Nutanix战略不清晰，到底走商用路线还是开源路线
  - 管理平台即有自己的，又要支持vCenter是因为vCenter和KVM不兼容问题

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

VMware VSAN，只是个零部件

vmware Virtual SAN

- 宏观角度
  - VMware VSAN只是个分布式存储软件，相当于HyperFlex的HXDP组件
  - 用VSAN搭建超融合，需要用户/代理商自己攒，最大问题是标准化很难，品质无法保障，后续服务很容易发生扯皮情况
  - 按容量收费，扩容成本高；必须企业版，成本高
- 产品和技术角度
  - 分布式存储技术中没有原生态的Dedup和压缩，不支持多Datastore
  - VSAN无法单独升级必须同时升级ESXi，会带来一系列升级问题
  - 网络NSX到现在也没整合进来，NSX和已有网络的兼容、性能、管理、QoS还任重而道远
  - 服务器和SAN技术服务都没有

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

DELL/EMC的VxRail

- DELL/EMC合并后有3种超融合，自己还没搞清楚未来方向
  - VxRail: 广达服务器+VSAN，原来EMC搞的
  - XC系列：OEM Nutanix，原来DELL搞的
  - VxRack：拿EMC ScaleIO+DELL服务器新攒的
- 估计XC系列，也就是OEM Nutanix的肯定是要死的
- VXRail原来是OEM广达服务器的，肯定是要改成DELL服务器的，转换过程品质和服务都难以保证
- ScaleIO是面向Server SAN的，不是面向超融合的
- EMC人不懂服务器，DELL人不懂存储，人员技术上咋合并还未知
- 合并完后还是没有网络人员，网络上的问题没能力管

© 2016 Cisco and/or its affiliates. All rights reserved. Cisco Confidential

基于开源的超融合主要是国内公司在搞，下面是比较

	Ciso HyperFlex 大公司，商用系统	华为 FusionCube 大公司，开源系统	H3C UIC 大公司，开源系统	深信服、天融信等 小公司，开源系统
技术路线	商用系统，自主研发	商用系统，开源软件	商用系统，开源软件	商用系统，开源软件 不管网络
产品	存储软件：自主研发 UCS服务器：自主研发 网络设备：自主研发	存储软件：Ceph开源 服务器：自主研发 网络设备：自主研发	存储软件：Ceph开源 服务器：OEM HP 网络设备：自主研发	存储软件：Ceph开源 服务器：OEM 网络设备：不管
服务	大公司，服务支持完善 网络、服务器、存储服务能力全面完整	商用软件和开源的兼容性，容易扯皮 开源碎片化问题，系统升级困难风险大	还不如华为	各部分服务能力都很弱
部署/运维	Cisco认证过整体标准配置，工厂预先标准化软硬件安装+代理商/客户应用部署；开箱即用快速上线，后期运维轻松简便 VMware服务器虚拟化市场普及度高，运维人员学习资料多学习成本低	整体有标准配置，部署安装必须厂家人员； Linux/KVM/Ceph难以运维，人力资源费用高； 开源系统的外包服务不靠谱，核心人员离职影响大； 商用软件和开源的兼容性问题是大问题。		
其它	SSD+高速SAS盘，磁盘高配	开源碎片化问题，用户容易被绑架		

Thank you.