

数据中心1314

硬件重构与软件定义



ZDNet 企业解决方案中心 首席分析师 张广彬
2014.1.10



数据中心行业趋势

- 互联网企业的规模需求与硬件重构
- 传统企业需要软件定义实现自动化

服务器 1,000,000 +

Google

amazon ?



Microsoft

服务器 100,000 +

amazon

Alibaba.com
阿里巴巴

YAHOO!

facebook

SAVVISSM

Tencent 腾讯

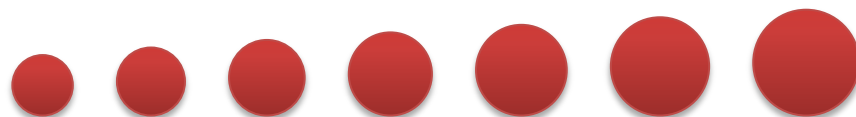
Akamai

Baidu 百度

SOFTLAYER[®]

rackspace[®]
HOSTING

服务器 50,000 +



从整机柜交付到模块化设计



Evolution of Rack Scale Infrastructure

Today

Infrastructure Disaggregation



- Shared Power
- Shared Cooling
- Rack Management

Next

Fabric Integration

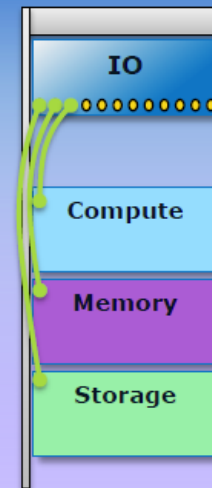


- Rack Fabric
- Optical Interconnects
- Modular refresh



Future

Fully Modular Resources



- Pooled compute
- Pooled Memory
- Pooled Storage

Enable flexible and efficient data centers through disaggregation of resources

整机柜交付-天蝎项目

FUJITSU



特点：

**共享电源、散热、管理
整机柜交付**

发起：百度、腾讯、阿里、中国电信

顾问：Intel

厂商：Dell、HP、IBM、华为、联想、浪潮、曙光

整机柜交付-Open Rack



发起：2011年发起，现由OCP基金会负责

成员：Facebook、Intel、Rackspace、Arista Networks、Goldman Sachs

包含项目：存储、主板与服务器设计、网络、机架、数据中心设计

Open Rack与天蝎



		天蝎V1.0	Open Rack V1.0
标准		19英寸机柜	改版自有标准
外部尺寸	高 (mm)	2100	2100
	宽 (mm)	600	600
	深 (mm)	1200	350-1220
内部尺寸	宽 (mm)	483	538
	高度单位	RU (Rack Unit)	OU (Open U)
	单位高度 (mm)	44.45 (1.75英寸)	48
	可用空间	44 RU	41 OU
IT设备	高度增量	1 RU	0.5 OU
	最小高度	1 RU	1 OU
	最大高度	20 RU	10 OU
共享设施	种类	供电、散热、管理	供电
	占用空间	4 RU	3×3 OU
网络设备	位置	中部	顶部 (ToR)
	高度	无额外占用	2 OU
设备供电		后端	后端
设备维护		前端	前端

OCP的模块化计算与存储

Storage



Server



冷存储-整机柜服务器



OCP冷存储方案

基本要求：最高容量，最低成本

参考规模：744个机架，24个接入，720个冷存储

冷存储机架配置：

2个双路计算节点

16个Open Vault

480个硬盘

每服务器接240个硬盘

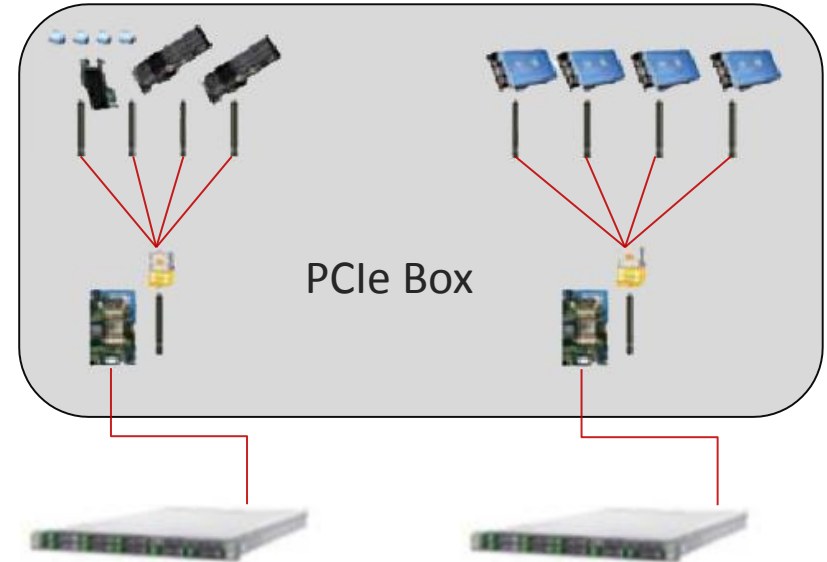
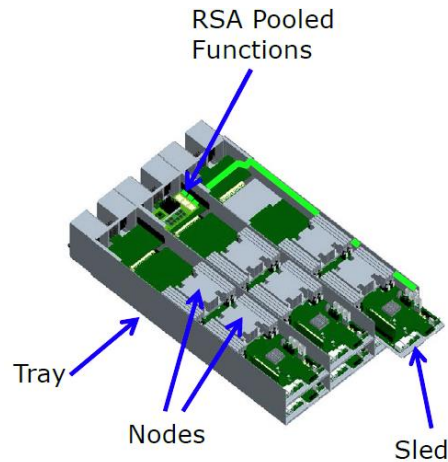
进一步精简机架、电源和网络配置

FACEBOOK的服务器类型



Standard	Type I	Type III	Type IV	Type V	Type VI	Type VII
Systems	Web	Database	Hadoop	Haystack	Feed	Cold Storage
CPU	High 2 x E5-2670	Med 2 x E5-2660	Med 2 x E5-2660	Low 1 x E5-2660	Med 2 x E5-2660	Med 2 x E5-2660
Memory	Low 16GB	High 144GB	Medium 64GB	Medium 96GB	High 144GB	High 144GB
Disk	Low 500GB	High IOPS 3.2TB Flash	High 15 x 4TB NL-SAS + 500GB SATA	High 30 x 4TB NL-SAS + 500GB SATA	Medium 2TB SATA	High Density 240 x 4TB SATA + 2TB SATA
Services	Web, Chat	Database	Hadoop	Photos, Video	Multifeed,	Cold Data
					Search, Ads	

机柜级服务器的愿景

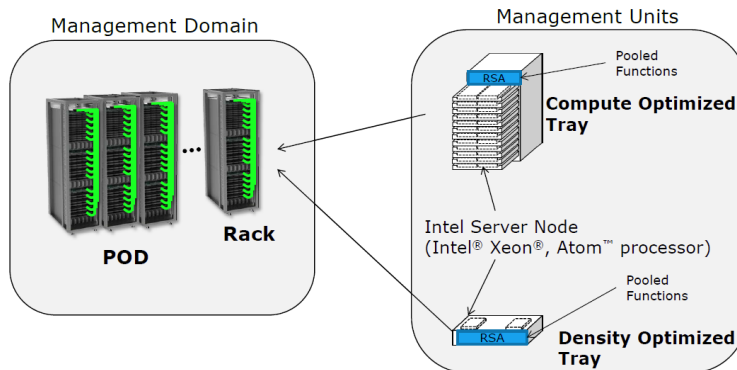


2 Fujitsu RX200 Servers

- 2 MXC connector with ClearCurve LX fiber
- 2 100G Intel® Silicon Photonic modules
- 2 Intel® Optical PCI Express Gasket ICs
- 2 Intel® Xeon® E5 2600 CPUs

PCI Expansion Box

- 2 Intel® Xeon Phi™ co-processors
- 2 100G Intel® Silicon Photonic modules
- 2 Intel® Optical PCI Express Gasket ICs
- 2 PCI Express switches
- 1 Raid Controller
- 2 SSDs



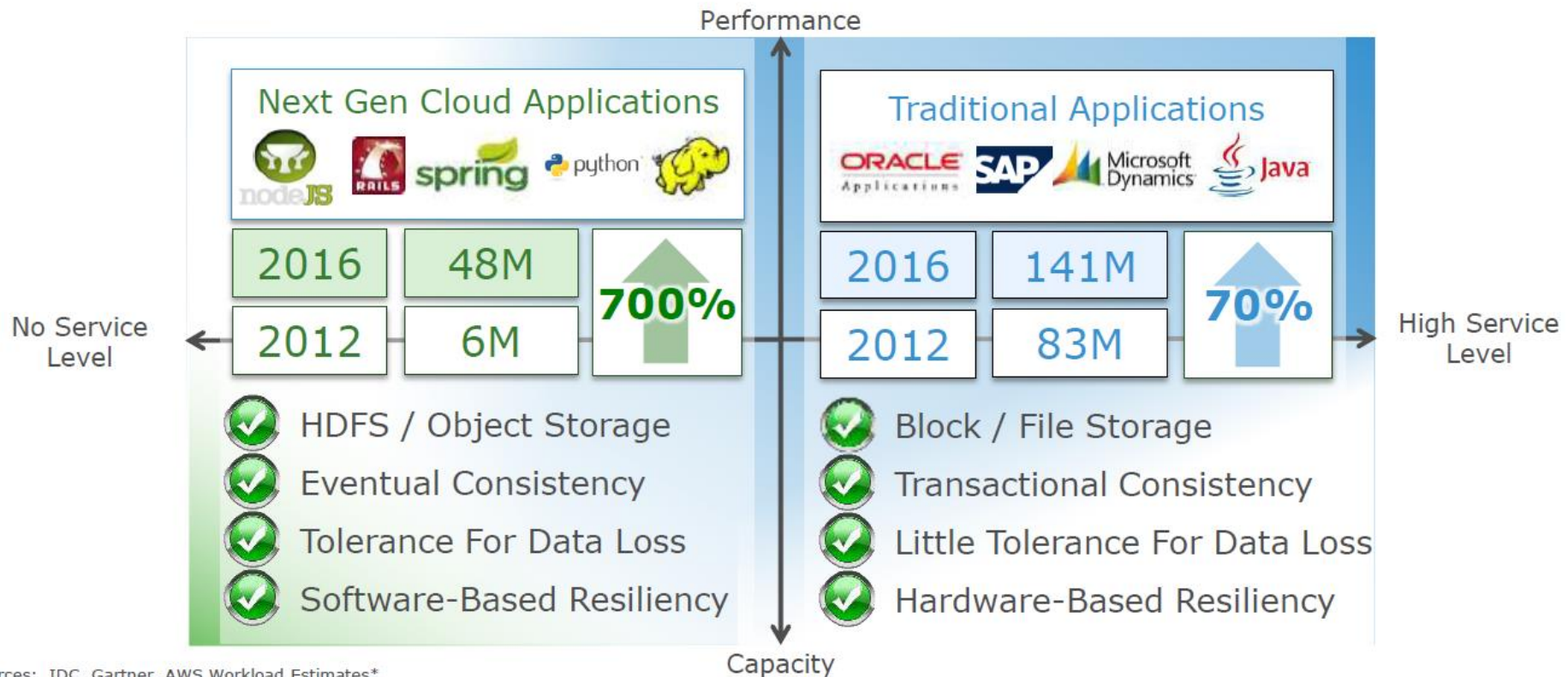
数据中心工作负载的多样化

Diversity of Datacenter Workloads



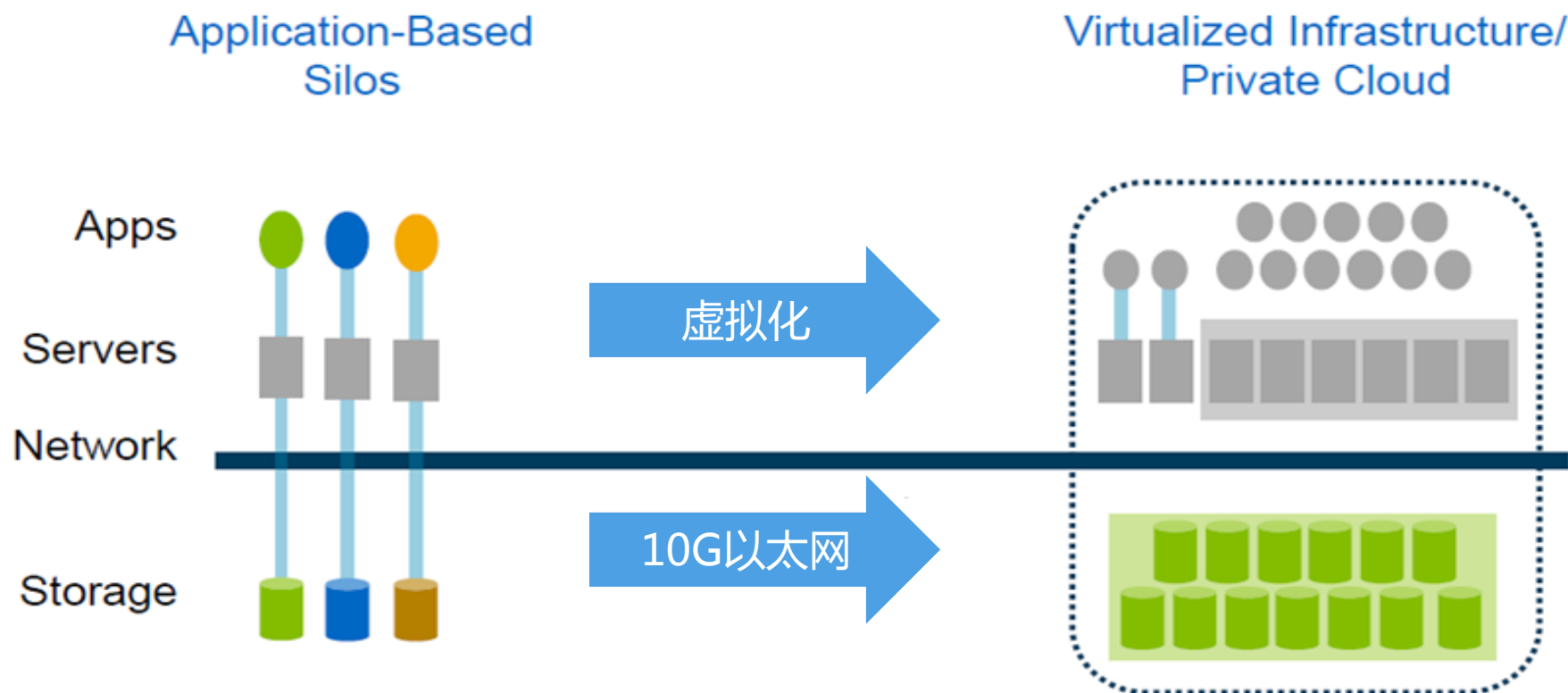
传统企业应用-规模不够 种类更多

2012-2016 Workload Growth



应用：互联网企业与传统企业

虚拟化成为模式分水岭





SDDC

软件定义的数据中心

软件定义数据中心与自动化



软件定义的数据中心

自动化

XaaS: 自助式目录、生命周期管理

vCAC

运营管理

基于分析的运营，实现自我修复、自我调整的 IT

vC Ops

Log
Insight

存储与可用性

Virtual
SAN

(公共测试版)

Virsto

计算

基于策略的控制和
基础架构服务自动化

vSphere

抽象化与池化

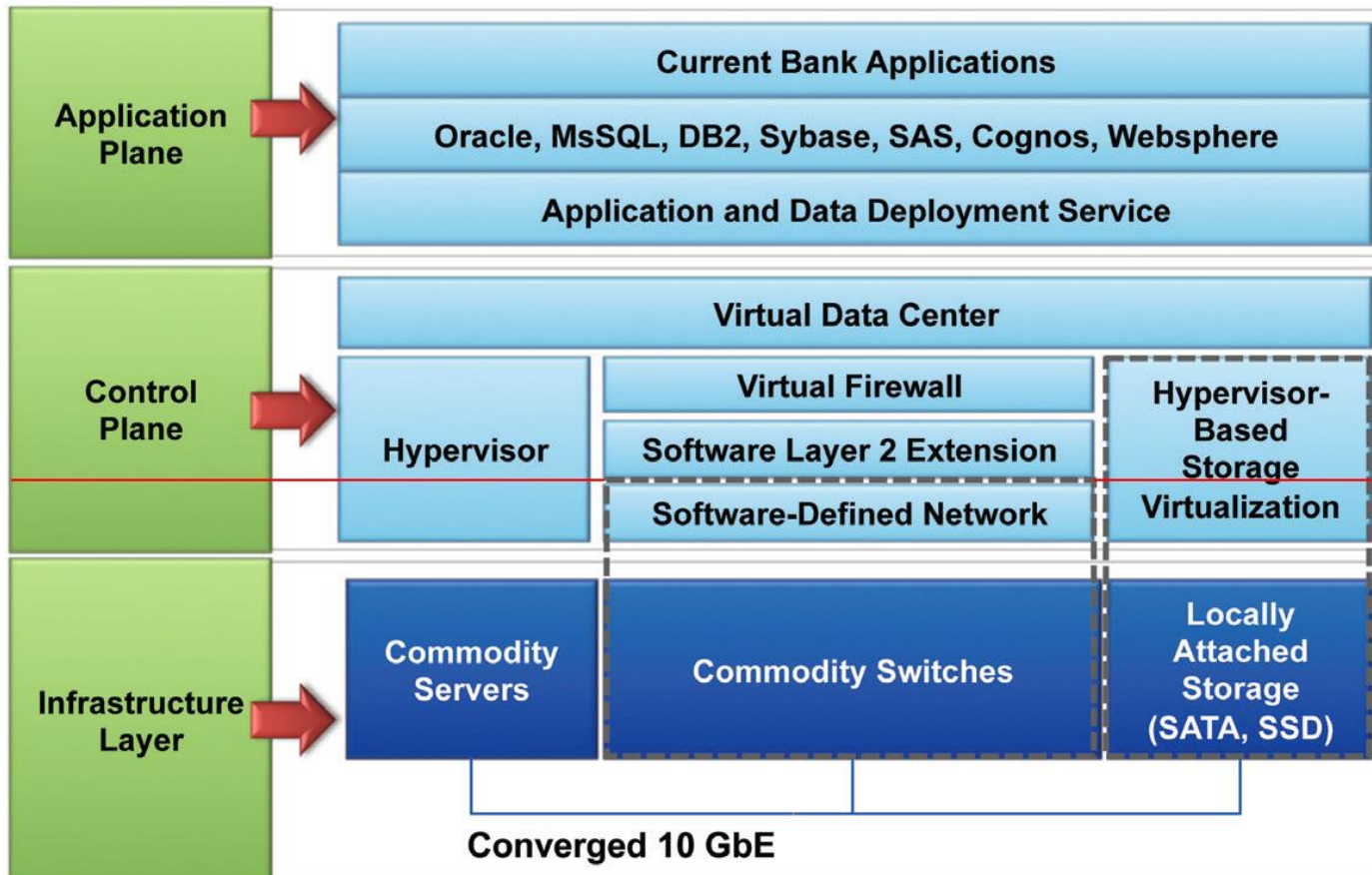
网络与安全性

NSX

软件定义数据中心构成



□ Key Feature ■ Hardware □ Software



Management Plane



范例：软件定义存储



软件定义特征-硬件无关性

vSphere HCL
上的服务器



← 1 Gb/10 Gb 网卡



← SAS/SATA 控制器 (RAID 控制器必须在
“passthru” 或 “HBA”
模式下使用)



← PCIe/SAS/SATA SSD

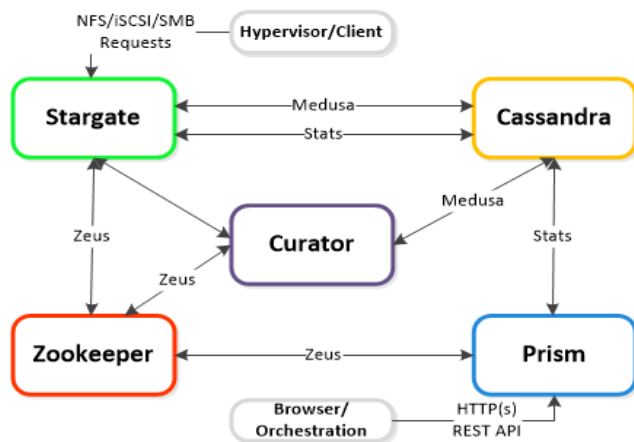


← SAS/SATA HDD

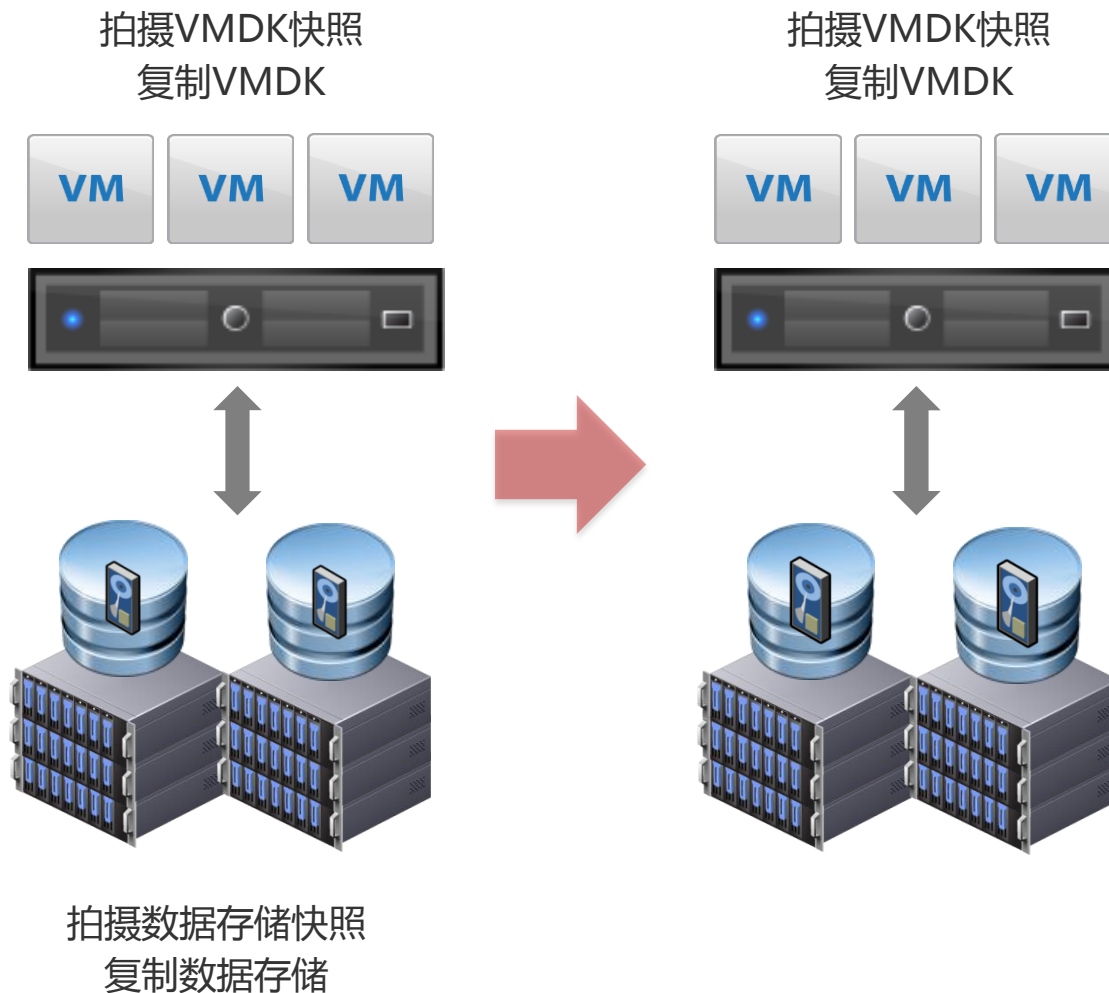
} 每服务器
至少 1 个

- 并不是 vSAN 集群中的每个节点都需要有本地存储
- 没有本地存储的主机仍然能够利用分布式数据存储

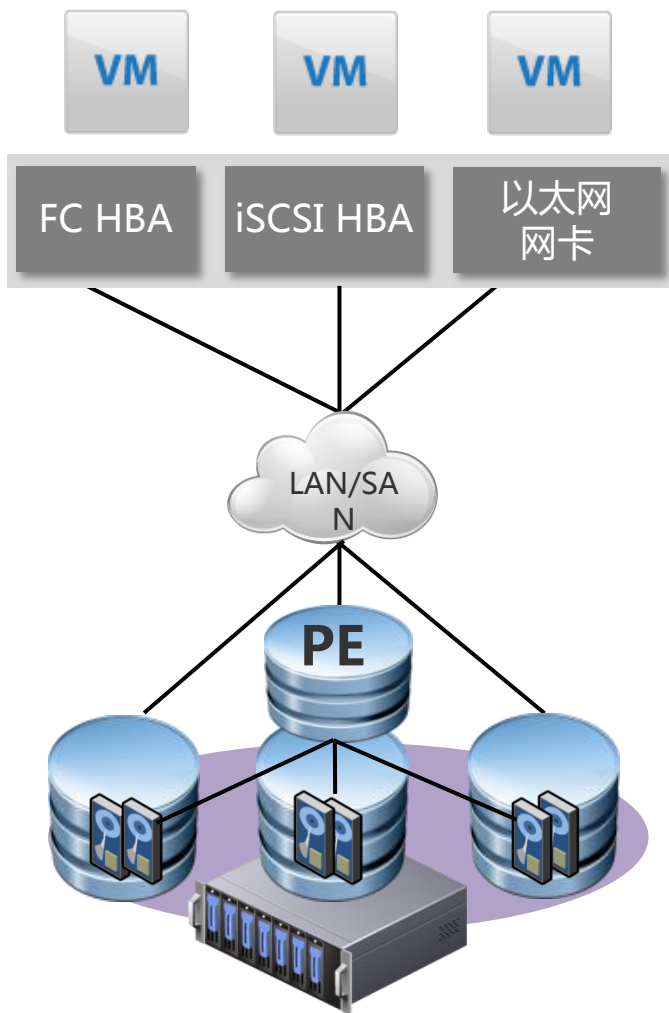
软件定义特征-大规模横向扩展



VVOL兼容传统存储系统



简化存储管理



1

无需配置LUN或NFS共享。

2

设置单一 IO访问实体[称为“协议端点” (PE)]以设置从虚拟机到 VVol的数据路径

3

设置逻辑实体[称为“存储容器” (SC)]将多个 VVol 分在一组，以简化管理



模式颠覆与行业重组

硬件自我颠覆

软件颠覆硬件

规模与自动化

传统巨头还没做好准备，您呢？

富士通中国论坛
Fujitsu Forum

2014.1.9-10

FUJITSU

shaping tomorrow with you

Reshaping ICT,
Reshaping Business and Society
重塑ICT 开启新未来

Thank you !