

---

# 并行发展的大数据基础架构



尹玉峰

戴尔解决方案资深经理

---

# 改变中的数据世界

4.3



成人拥有的平均数字设备

85%



来自新的数据类型

10X



每五年增长

27%



使用社交媒体

至2015, 使用现代信息管理手段的组织, 在财务表现上超越同行20%

"Information Management in the 21<sup>st</sup> Century"



# 数据以更快的速度产生与消费

**\$232**

billion dollars 到2016,  
预算在大数据的投入

**70%**

的数据来自消费者。但是企业负责存储  
并管理器中的80%

**247 billion**

Emails/Day  
80% 垃圾邮件

**4.4 million**

大数据创建IT职位  
只能提供1/3

**\$600 billion**

每年因为低质量数据导致的浪费

**48 hours**

每分钟上传到Youtube的  
视频

**37.5%**

大型机构承认大数据分析  
是他们最大的挑战

**1.8 Zettabytes**

of business data in use in 2011

**200 million**

每天上传到Facebook图片数



# 大数据

## 信息增长.....

### Volume

A large amount of data, growing at rapid rates

### Variety

Wide range of data types and structure

### Velocity

Data that must be processed at high speed to facilitate rapid decisions

导致数据太大或太复杂，传统数据管理工具难以管理



捕获



管理



分析



# 大数据中的基础架构

## 有这样的基础架构



# 大数据中的基础架构

## 还有这样的基础架构



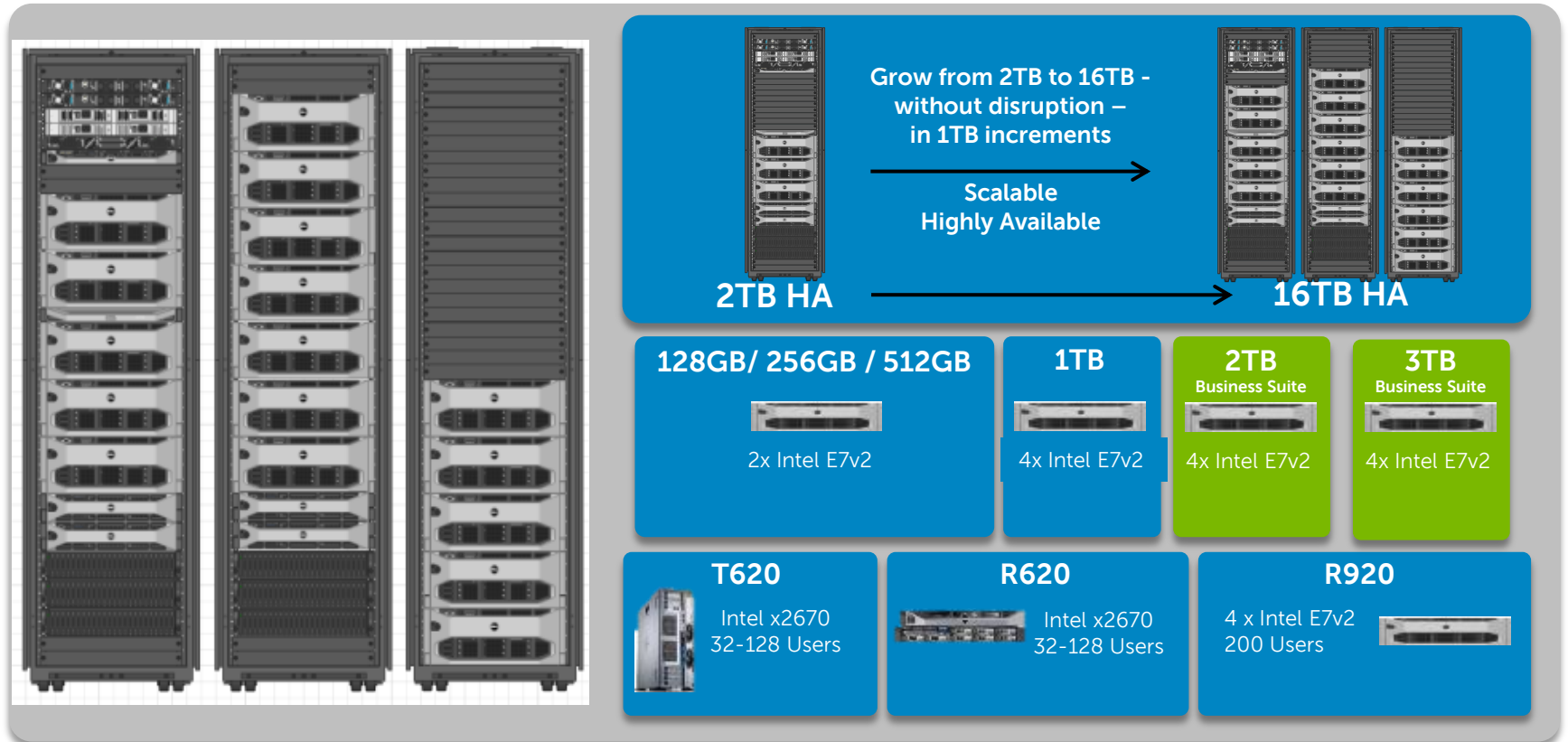
# 大数据中的基础架构

也有这样的基础架构



# SAP完整认证的HANA方案

还有一个很重要的SAP

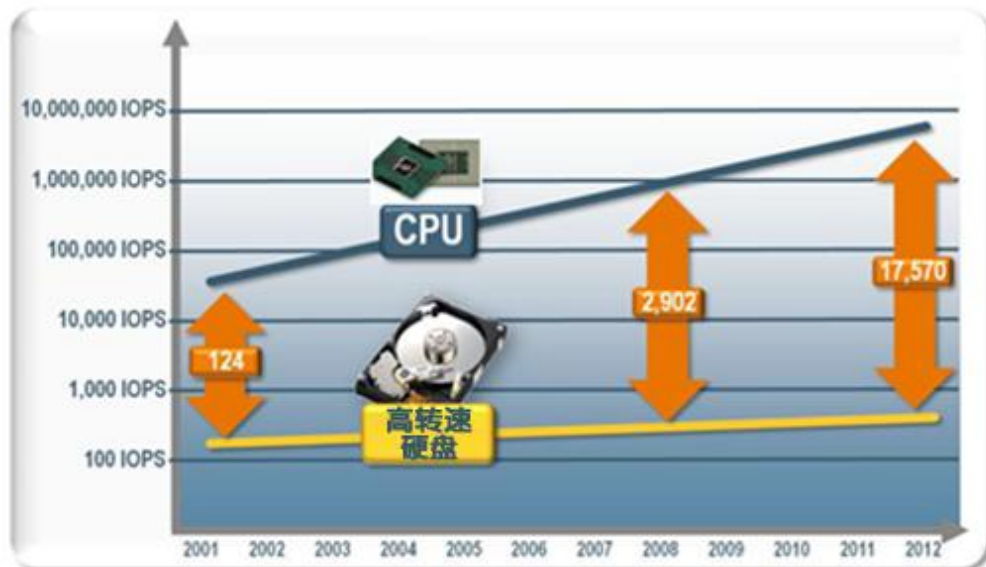




# 客户面临的难题： 应用程序的性能受到限制

*CPU、内存和网络性能不断提升，同时存储容量也出现了增长……但应用的速度并未跟上步伐。*

## 传统驱动器的问题



- CPU能够达到将近1000万IOPS，而最出色的15K SAS驱动器也只能达到540 IOPS左右。
- 采用内存来克服存储I/O瓶颈的做法成本高昂。
- 当事务处理速度十分缓慢时，原因往往是存储层出现磁盘争用，从而增加I/O延迟，进而减慢事务处理速度。

# 闪存领域创新

## 高密度1.8英寸固态硬盘和2.5英寸NVMe PCIe固态硬盘

性能优异、经久耐用、敏捷灵活



- 采用针对固态硬盘进行了优化的服务器配置，从而可加快数据库性能
  - OLTP工作负载
  - 联机分析
  - CRM



- 提供无可匹敌的性能——高IOPS、超低延迟
- 是Dell Fluid Cache for SAN和新的SanDisk DAS Cache的组成部分
- 可从正面检修、支持热插拔，从而简化了管理



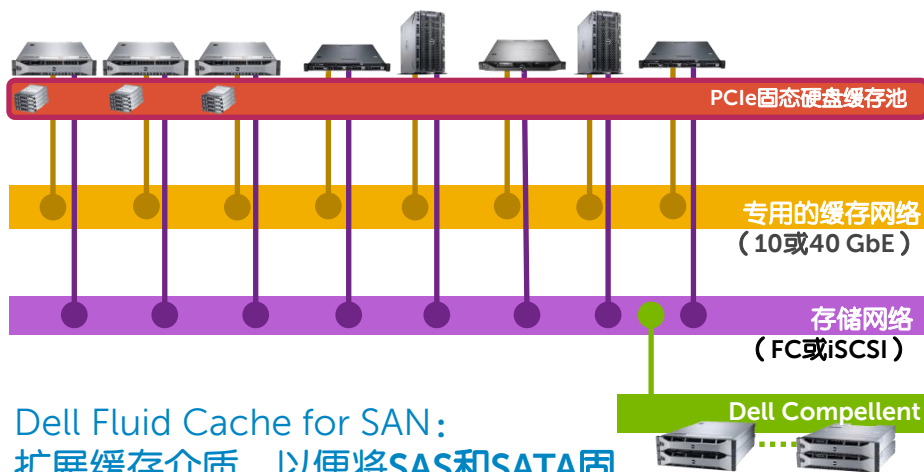
# PowerEdge解决方案：加快应用程序性能

## 闪存存储和缓存选项

1.8英寸SAS和SATA固态硬盘，以及热插拔的PowerEdge Express Flash NVMe PCIe固态硬盘。

### Dell Fluid Cache for SAN

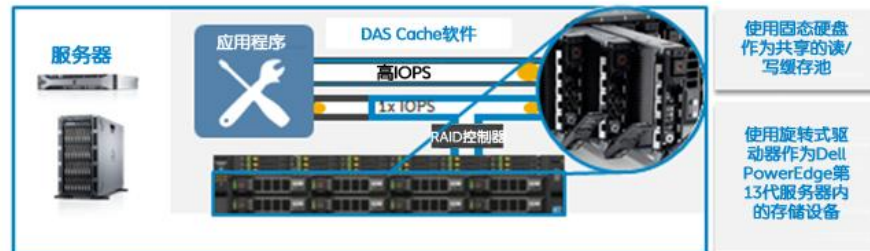
- 为数据库和虚拟化工作负载提供非凡的应用程序响应速度。
- 可通过添加Dell PowerEdge Express Flash NVMe PCIe固态硬盘或服务器节点来动态地扩展性能——甚至是在实时缓存池内也是如此。
- 利用回写式高速缓存技术以及与Compellent后端存储的完全一致性来确保数据完整性。



Dell Fluid Cache for SAN:  
扩展缓存介质，以便将SAS和SATA固态硬盘也纳入为新一代服务器的缓存！

### SanDisk DAS Cache

- 这是戴尔推出的服务器级DAS（直连存储）I/O应用程序加速解决方案，旨在与新一代PowerEdge服务器搭配使用
- 可使OLTP、OLAP和业务分析等工作负载发挥出超凡性能。
- 既支持Linux操作系统，也支持Windows操作系统
- 用途广泛，支持使用各种各样的设备作为缓存：SATA、SAS和PCIe固态硬盘；PowerEdge Express Flash NVMe PCIe固态硬盘；以及基于PCIe的扩展卡，例如Fusion IO。
- 旋转式存储设备位于服务器内部，同时采用精选的Dell PowerVault存储设备作为外部存储。



# 容量方面的增强

容量更大、速度更快、安全性更高

*总体系统存储优化扩展至组件级别*



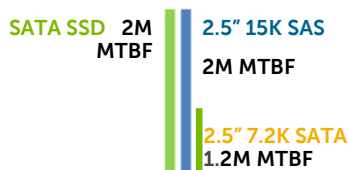
- **容量更大**

- 容量优化型3.5英寸驱动器将从6TB升级为8TB
- 容量优化型2.5英寸驱动器的容量将达到2TB
- 性能优化型驱动器的容量将达到1.8TB

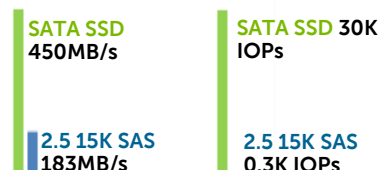


# SSD vs HDD

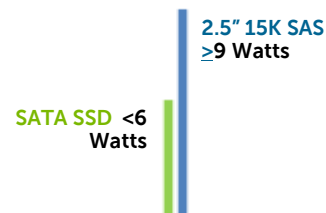
- 提供与15K SSD 接近的入门价格，但是**100X** 性能提升
- 相同可靠性 vs 15k rpm SAS HDD
  - 2百万小时无故障时间
- 50%省电 vs 15k RPM SAS HDD
  - SSD 6W vs HDD 9W
- 超快速的Raid Rebuild时间，提高uptime时间



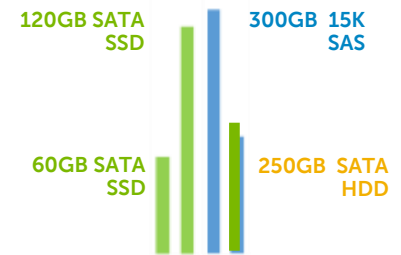
可靠性对比



性能对比



能耗效率



Entry Price Point Replacement

# 满足需要提高数据I/O能力的应用程序的需求

PowerEdge R630机架式服务器

**配备24个1.8英寸固态硬盘的  
R630可使本地存储的性能提升  
240%，同时使成本保持大致不  
变**



- 对85%至90%的Oracle RAC环境而言，R630都是理想选择
- 由于消除了数据I/O瓶颈，因此服务器的利用率更高。
- 可以采用后端存储来提供容量而非性能。
- 实现了机架密度的翻倍，同时在1U密度既定的情况下还降低了功耗、冷却和空间成本。



# 集闪存、高速缓存和高容量存储于一体

R730xd完美搭配了所有这三个方面，同时仅占2U空间

用户需要兼具闪存性能与庞大本地存储容量的存储解决方案，以便为大数据(Hadoop)、Microsoft SQL数据库和数据集市等工作负载提供强劲动力。

## 以一敌二

- 原先需要两台机器，一个做密集读写，一个存储大容量数据
- 现在可以通过一台混合18\*1.8"SSD+8\*3.5"HDD的R730xd实现



- Hadoop针对I/O密集型Map-Reduce应用程序使用闪存和大容量存储。闪存设备的性能超过传统磁盘3倍

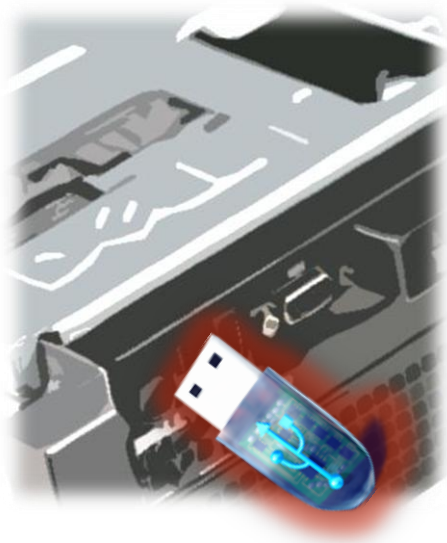
配备18个1.8英寸固态硬盘和8个3.5英寸普通硬盘的PowerEdge R730xd机架式服务器



# 通过U盘使用iDRAC Direct

使用iDRAC Direct轻松完成配置

*如果您就在系统旁，可以使用U盘轻松完成服务器设置，  
而无需使用线缆和笔记本电脑*



XML  
配置文件

- 使用方便的标准U盘即可将新配置安装到您的戴尔第13代服务器上
- 只需不到1秒时间即可将BIOS、iDRAC、PERC和网卡的配置上传到iDRAC



# 通过USB线缆和笔记本电脑使用iDRAC Direct

通过本地笔记本电脑连接到iDRAC GUI

*从此告别应急方案！可以在现场直接、方便地建立连接来执行高级管理任务。*



- 可通过USB线缆连接到笔记本电脑来访问iDRAC GUI
- 可以直接与iDRAC GUI进行交互来进行高级服务器管理和维护（即使尚未配置iDRAC联网设备，也是如此）
- 可以检查详细的日志文件以便加快服务器修补速度
- 正面的USB端口插拔起来十分方便
- 强大的iDRAC凭据以确保安全访问  
无需专用的软件或驱动程序！

# iDRAC快速同步

使用您的智能设备通过NFC连接到iDRAC

*借助iDRAC快速同步，使用您的移动智能设备即可获取和设置关键的服务器信息*



- 只需清扫您的手机即可快速获取关键的服务器信息和错误日志
- 还可以在几秒钟的时间内配置好关键设置
- 此功能随R630、R730和R730XD服务器提供
- 此功能会用到OpenManage Mobile应用程序，并且可以选择使用支持

**业界首款提供集成式近场通信技术(NFC)支持的服务器!**

# 经过优化的工作负载配置文件

采用经过预先测试的BIOS设置，实现优化的性能

*快速轻松地为您的工作负载获取正确的设置*



- 借助新的BIOS设置，您可以快速选择一种经过预先测试、可使特定工作负载性能达到最优的服务器配置
- 初期提供可使HPC和高频交易(HFT)工作负载发挥最佳性能的配置文件（以后将提供更多配置文件）
- 与标准服务器配置相比，可使性能提升XX%

*节省为实现最优性能而查找最佳设置所耗费的时间和工作量！*

# 加快服务器的启动速度

启动速度加快60%

*最大限度减少服务器停机时间并加快动态配置速度*



- 显著缩短诸如固件更新、操作系统修补或断电重启等事件造成的停机时间
- 操作系统修补和固件更新通常需要重启服务器；缩短启动用时可以加快这些过程的速度
- 可以更快地启动生命周期控制器 GUI

# 设想一下，如果有一种IT方法...

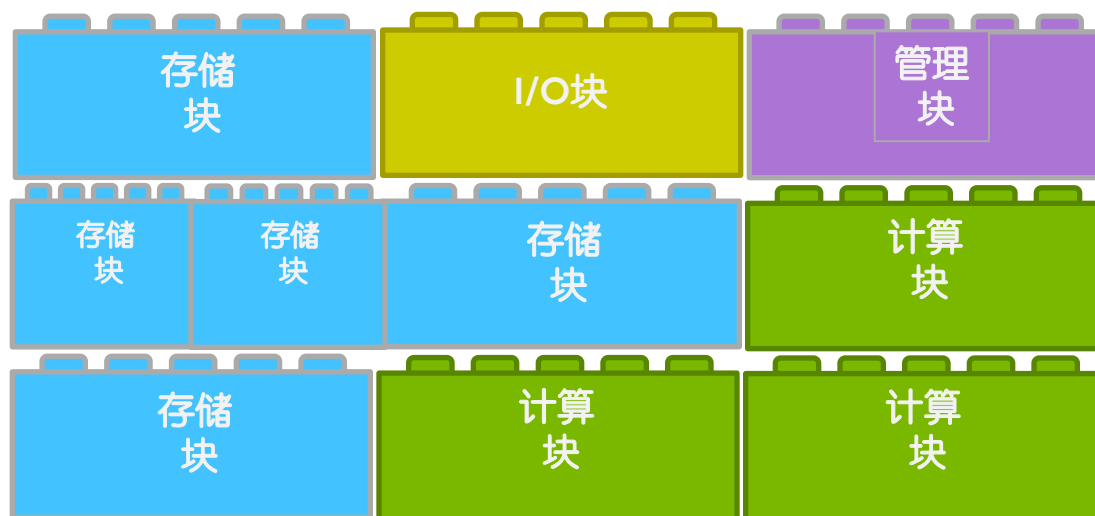
可以使得构建某种工作负载解决方案变得像儿童搭积木一样简单.....



如果您可以像搭小块积木那样循序渐进地配置IT资源，将会如何？

# 试想，如果您能做到以下几点，那会如何...

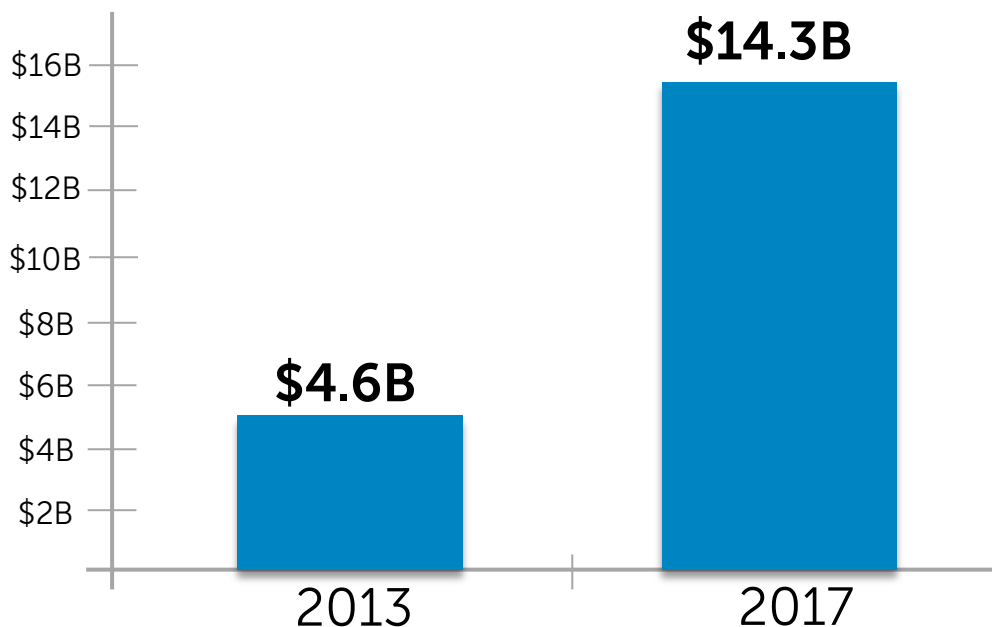
...完全按照您需要的资源种类、数量和时机配置资源



...让您的基础架构恰到好处地满足您的工作负载需求，而不需过度配置

# 持续增长的融合系统

全球集成系统市场预测 **30%**年度复合增长



\* Source: IDC



# PowerEdge FX: 灵活与适用

## PowerEdge FX2



2U 共享基础架构

- 77% 节省看空间
- 1/3 节省电力
- 80% 电缆连接减少

## 运算



PowerEdge FC630

## 高密度虚拟化



PowerEdge FC430

## Web服务



PowerEdge FM120x4

## 数据库



PowerEdge FC830

## 灵活 DAS



PowerEdge FD332

## IO 聚合器



PowerEdge FN IOA





# Dell 存储

## SC4020



全闪存阵列

72% 降低全闪存成本

**Enterprise storage advantages for small and mid-sized deployments**

### Enterprise performance

Workloads run faster with AFA or hybrid solutions; put ALL your hot data on flash

### Best-in-class intelligence

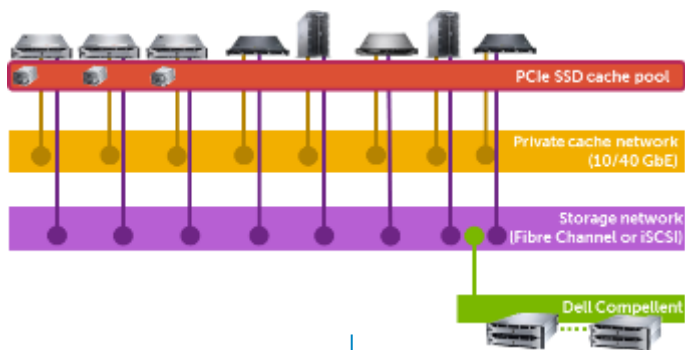
Self-optimizing, self-provisioning, auto-tunes your datacenter

### Unprecedented value

Ideal for SME standalone or ROBO deployment with SC8000



# Fluid Cache for SAN – 经济的方式把闪存性能发挥到极致



>5 百万  
IOPS

3.9x 倍交  
易数提高

**Faster response times,  
unprecedented I/O performance for  
applications such as OLTP and VDI**

## **Attain greater performance**

Deliver results faster for business critical apps and workloads

## **Achieve enhanced flexibility**

On demand acceleration without over-provisioning

## **Realize maximum stability**

Safeguard data from server to SAN with no single point of failure

# 全面开放SDN

Standard orchestration & automation tools

Optional 3<sup>rd</sup> party SDN

Any networking OS

Open standard hardware

Merchant silicon



Automation



Orchestration



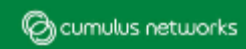
Monitoring



vmware



Others



Others



Others




Others

# 融合无妥协

Most Likely  
Converged Infrastructure  
to be Recommended by  
**Customers**<sup>1</sup>



**2.3x** more  
compute nodes  
per rack



**6x** faster implementation  
of new virtual  
infrastructure



**1** console for  
deployment **8**  
instead of



**99%** fewer steps  
to deploy a complex  
virtualization  
infrastructure



**45%**  
better system  
performance/watt



<sup>1</sup>Technology Business Research, Inc. "Converged Infrastructure Landscape Report" by Christian Perry, Krista Macomber, Angela Lambert, Matt Bowden; June 2013.

