

HUAWEI ENTERPRISE ICT SOLUTIONS **A BETTER WAY**

# 华为Tecal ES3000 PCIe SSD

## 高性能存储卡 主打胶片

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

HUAWEI TECHNOLOGIES CO., LTD.



# 目录

- PCIe SSD存储卡将成为SSD的主流
- Tecal ES3000规格及亮点
- Tecal ES3000应用场景及应用效果
- 友商PCIe SSD产品对比分析

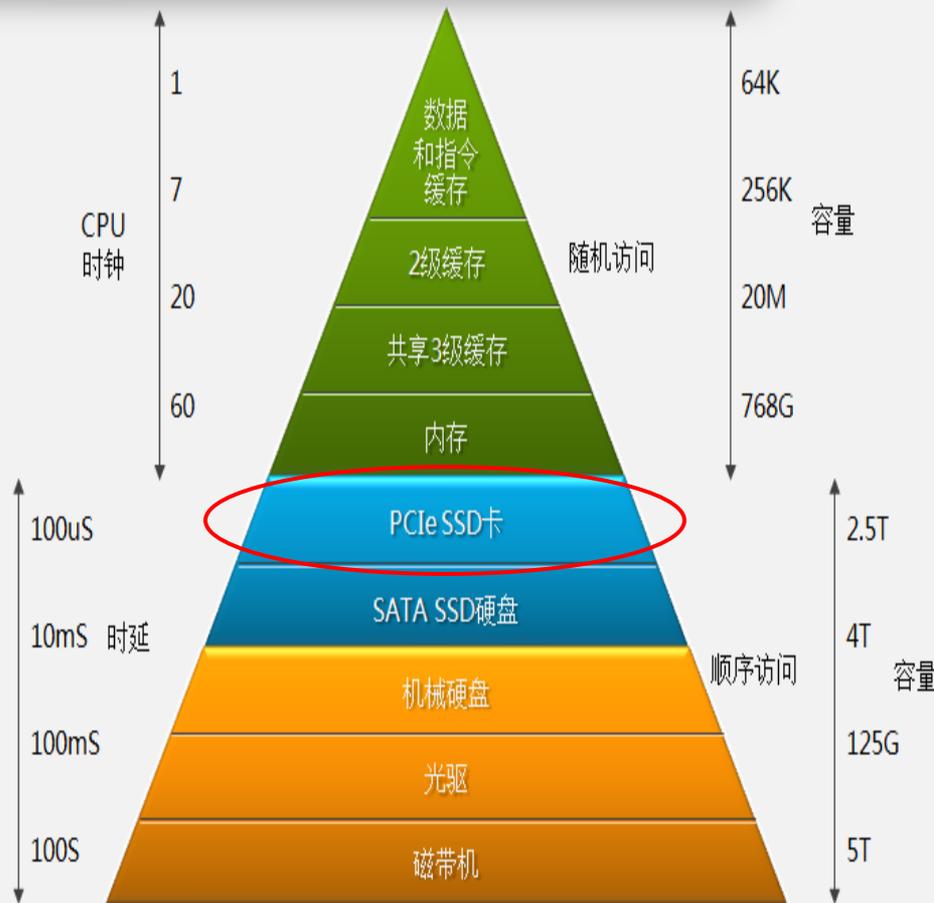
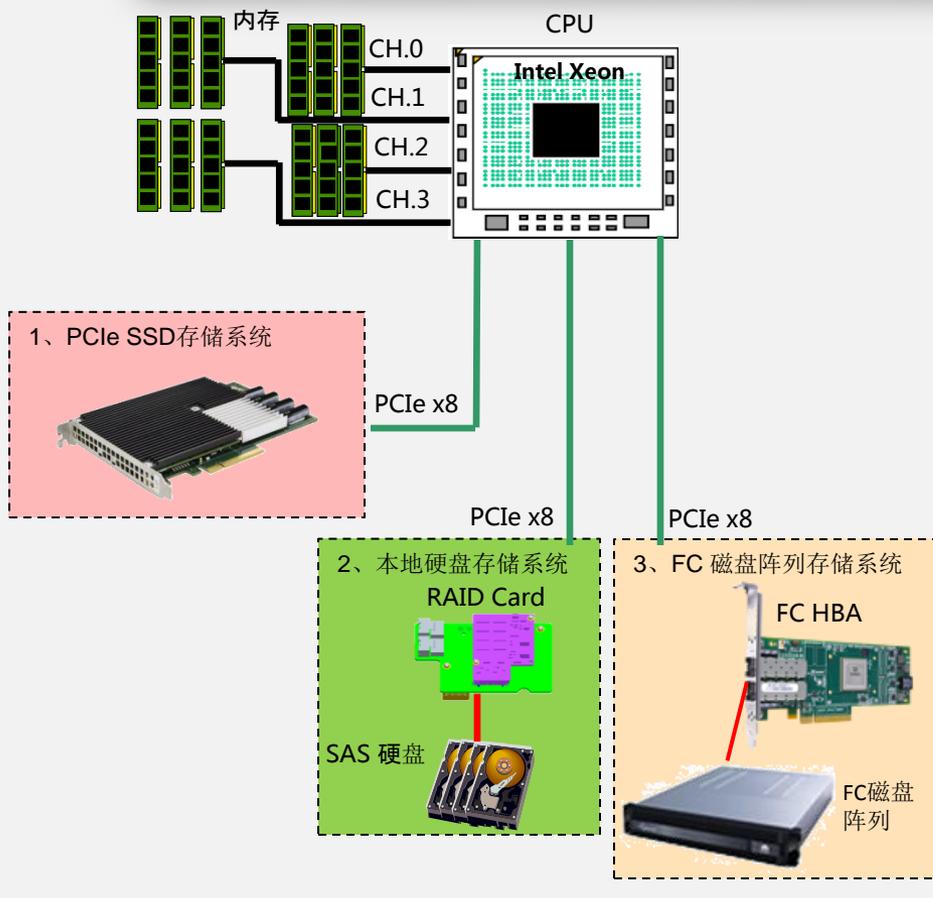
# CPU与硬盘性能差距加大，IO瓶颈问题日益突出

用户和应用的几何级增长，处理器与存储性能的极度不平衡发展，导致I/O失衡现象越发明显，只有加速数据读写性能才能维持业务稳定运行，提升用户体验效果！



# PCIe SSD简化存储路径，解决存储IO瓶颈

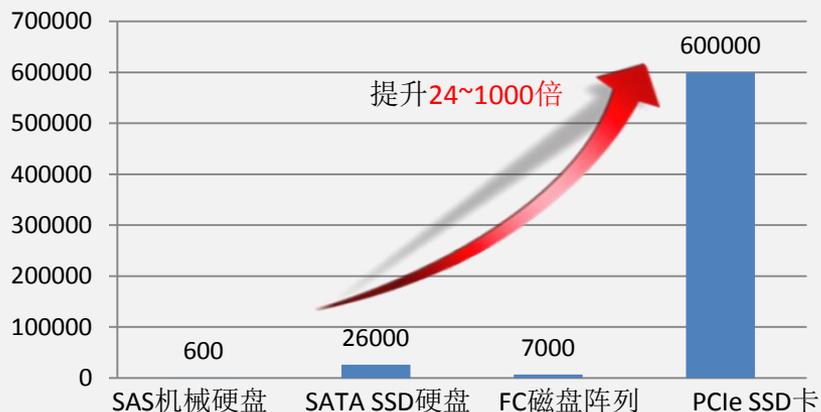
PCIe SSD直接与CPU通信，缩短存储访问路径，结合flash高速的读写性能，极大的提升存储性能，解决存储IO瓶颈。



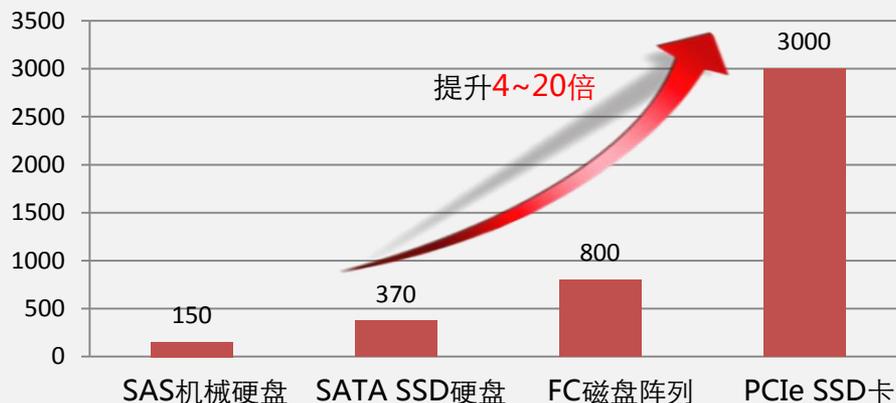
# PCIe SSD卓越的读写IOPS、带宽性能

PCIe SSD提供高达600K@4KB 100% Random读IOPS性能，以及3000Mbps读写带宽，满足客户对高IO、高带宽的业务需求。

## 100%随机读IOPS @4KB



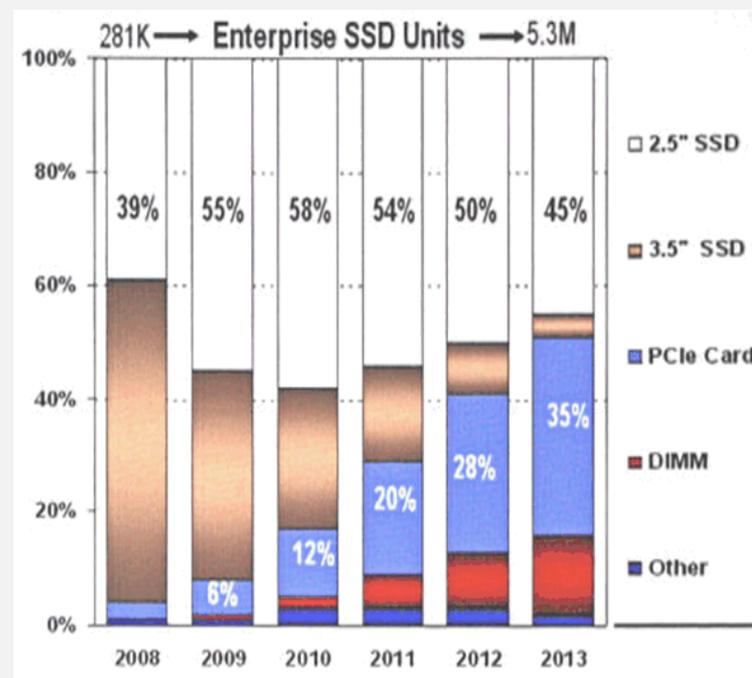
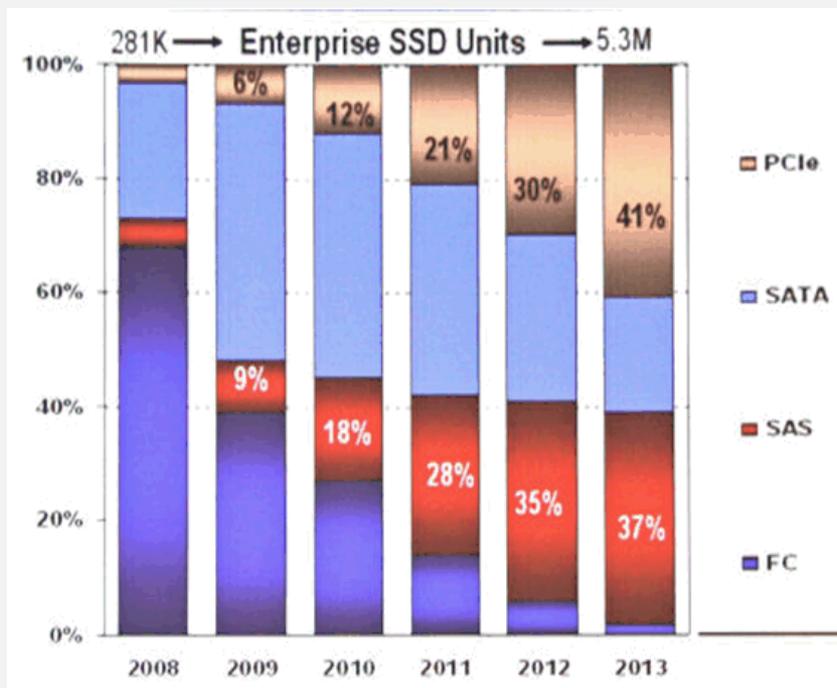
## 最大读写带宽 (Mbps)



	本地硬盘存储系统		FC 磁盘阵列存储系统	PCIe SSD系统
单体最大容量	600GB(SAS)	400GB(SSD)	7.2TB(12个SAS盘)	2.4TB
单体100%随机读IOPS	0.6K(@4KB)	26K(@4KB)	7K(@4KB)	600K(@4KB)
单体最大带宽	150MB/s	370MB/s	800MB/s	3000MB/s
平均响应延时	10~100 ms	0.1~1ms	1~10ms	10~100 μs
每GB IOPS性能	低	较高	低	最高

# PCIe SSD存储卡将成为SSD的主流

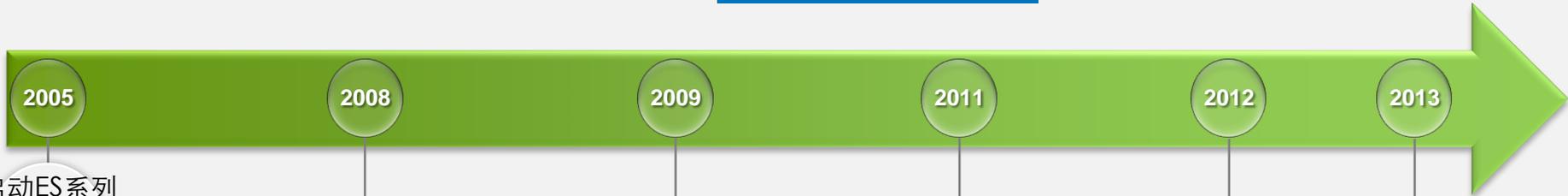
PCIe SSD凭借其性能和成本优势，已经成为SSD的新宠，并且快速上量，在未来几年中将成为主流高性能大容量存储介质



数据来源于Gartner

# 华为7年PCIe SSD产品经验

市场客户



启动ES系列  
PCIe SSD  
产品项目

**第一代**  
128GB/256GB  
SLC

**第二代**  
50/34nm  
512GB/1TB  
MLC

**第三代**  
ES2000,  
34/25nm  
640GB, MLC

**第四代**  
ES3000, 25nm  
0.8/1.2/2.4TB  
MLC

**第五代**  
20nm  
规划中

产品



# ES3000 产品规格

型号	ES3000		
容量	800GB	1.2TB	2.4TB
Flash颗粒	MLC	MLC	MLC
最大读带宽	2.2 GB/s	3.2 GB/s	3.2 GB/s
读IOPS (4KB, 100%随机)	570K	760K	760K
读延时	49μs	49μs	49μs
最大写带宽	1.2 GB/s	1.8 GB/s	2.8 GB/s
写IOPS (4KB, 100%随机)	120K	180K	240K
写延时	8μs	8μs	8μs
混合读写IOPS (4KB, R/W:7/3)	260K (R/W: 180K/80K)	400K (R/W: 280K/120K)	430K (R/W: 300K/130K)
功耗	20~35W	25~50W	25~60W
重量	300g	350g	350g
总线接口	PCI-Express 2.0 x8		
尺寸 (长×宽×高)	全高半长		

# 高性能、高可靠、易管理、绿色节能

- 大容量：800GB/1.2TB/2.4TB
- 高IOPS：随机读760K@4KB，随机写240K@4KB
- 高带宽：PCIe2.0 x8接口，带宽高达3200MBps
- 低访问延迟：8 $\mu$ s

- PCIe接口，直接与CPU通信
- 动态/静态的磨损平衡算法
- 内置FTL、RAID5、ECC、数据巡检等引擎，减少主机资源消耗

持续的高速  
读写性能

内置优化的  
算法引擎

多重的  
数据保护

超低功耗  
绿色节能

灵活的配  
置管理



ES3000

- RAID5、ECC、数据巡检等技术
- 通道间的冗余共享（单通道失效）
- 支持TRIM技术高效回收闲置Flash数据块
- 智能的冷热数据分类
- 超级电容掉电保护

- 支持主流OS及虚拟化系统
- 丰富信息查询管理工具
- 提供SNMP查询接口，方便统一网管
- 灵活设置可用容量，实现容量和性能的最佳组合

- 电子存储芯片，超低功耗
- 非机械结构，无噪音污染，寿命长
- 节省数据密集型服务器群组的电能和冷却成本

# 可靠性铁三角：全方位、全时段保护客户数据

## 实时保障

- 动态以及静态的**磨损平衡算法**
- **ECC纠错**能力达到30bits/1KBytes，超强纠错能力，确保写入数据成功
- TRIM技术高效回收闲置Flash数据块，确保性能保持最优

## 可靠性铁三角

- 内置动态**RAID5**避免单颗Flash损坏导致数据丢失
- 基于超级电容的**掉电保护**，确保数据在异常掉电情况数据仍可完整保存

## 提前预防

- 定期**巡检数据**，搬移出错数据，预防数据出错，确保错误不累积
- **通道间冗余、资源共享**，确保在多个芯片故障下能正常工作
- **数据分类**技术减少数据写放大，提升Flash擦写寿命

## 事后补救

# 提升客户业务价值

## 降低存储延迟，消除I/O瓶颈

- 提升应用性能，改善用户体验

## 处理最有价值的数

- 把访问频繁、最有价值的数  
据存储在ES3000上，将满足  
超过80%的IOPS需求，达到  
商业价值最大化



## 应用加速，缩短响应时间，提升效率

- 大幅缩短响应延迟，明显提  
高应用响应速度，实现系统  
长期高效运行

## 优化IT基础

- 减少服务器及存储的部署数  
量，节省部署数据密集型服  
务器群组的电能和冷却成本

# 丰富的应用场景及业务特点



**业务特点：**主要对1KB~100KB的小文件进行高频度的读取操作，追求高IOPS。

## Web搜索/通行证(登录)业务

更快的结果返回



## 高性能计算

更短的计算周期

**业务特点：**要求在较短的时间内完成高并发的海量数据处理，需要很快得出计算结果，为用户的决策做参考。

# 高性能存储



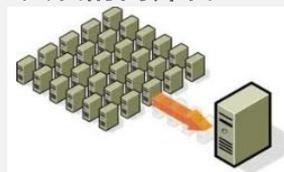
## 数据库应用

更快的查询速度

**业务特点：**高并发，结构化数据量大，目标数据块小且离散，要求存储容量大、IOPS性能高。

## 虚拟化应用

突发的高并发IO



**业务特点：**虚拟机登陆高峰期时产生突发的高并发IO需求，要求高IOPS性能。

## Email/ERP办公自动化

更高的办公效率



**业务特点：**高并发、高IOPS、低延迟，对存储介质随机读写性能要求较高。



## 大数据/Cache/网络加速

更快的热数据访问速度

## 视频点播/资源下载

更多的用户并发点播



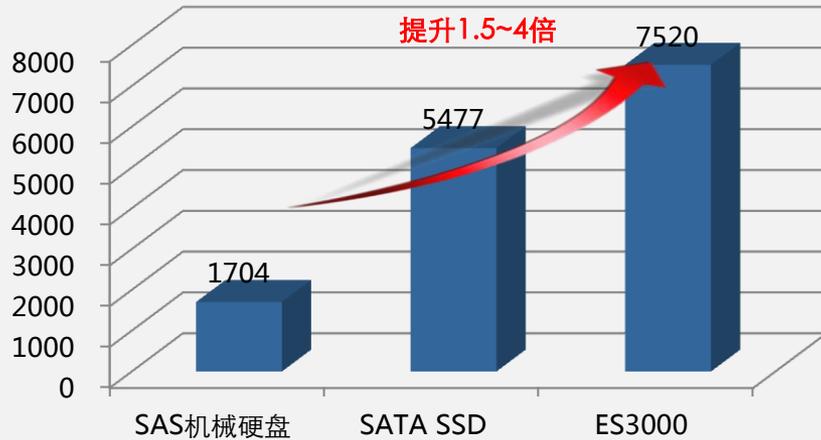
**业务特点：**存储文件较大（一般>10MB），文件更新频度很小，读多写少，要求存储高带宽。

**业务特点：**高并发，非结构化数据量大，目标数据块小且离散，将经常访问的数据（热数据）存放在SSD上，可以很好的实现加速。

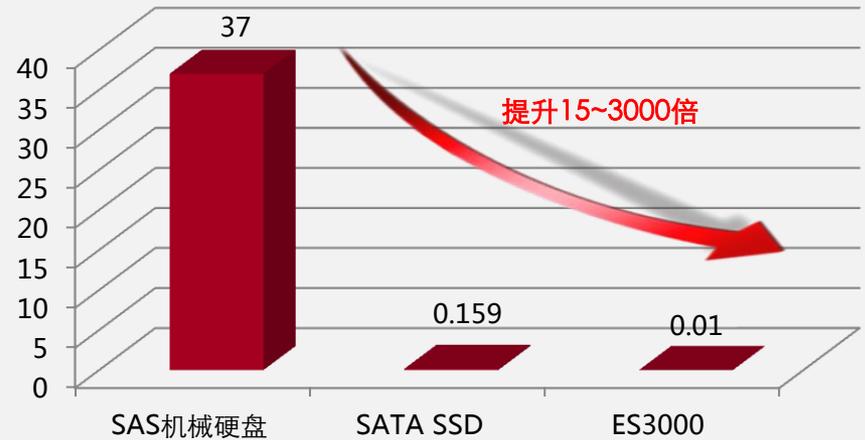
# SQL Server 2012查询性能提升4倍以上



每秒查询性能



平均查询时间(ms)



Fast Track for SQL Server 2012							
	服务器配置描述	最大用户数据容量(TB)	额定用户数据容量(TB)	平均查询块大小(KB)	每秒查询性能	查询I/O性能 (MB/s)	平均查询时间(ms)
SAS机械硬盘	RH2288 V2: 2*E5-2690, 16*8GB, 12*600GB SAS 机械硬盘	8	5.5	436	1704	743	37
SATA SSD	RH2288 V2: 2*E5-2690, 16*8GB, 12*400GB SATA SSD	7	5	438	5477	2399	0.159
ES3000	RH2288 V2: 2*E5-2690, 16*8GB, 2*2.4TB ES3000	10	8	427	7520	3211	0.01

# 用友U8 ERP应用性能提升30%以上

用友ERP-U8

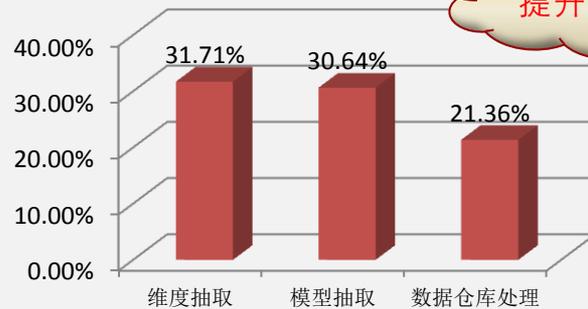
## 并发应用效果提升

平均效率提升20%



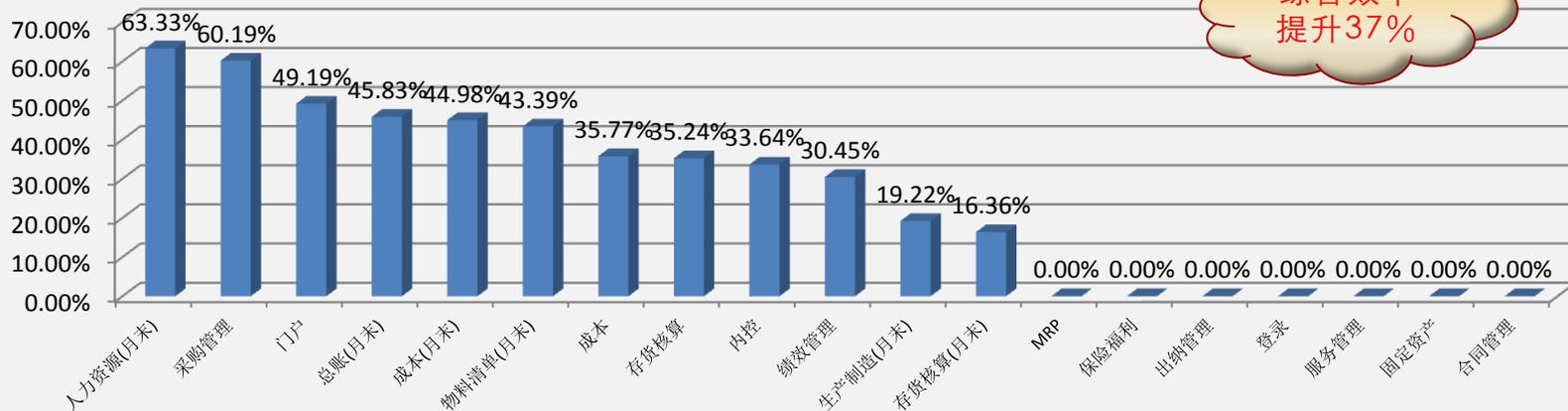
## BI场景效率提升

平均效率提升30%



## 月末处理场景效率提升

综合效率提升37%

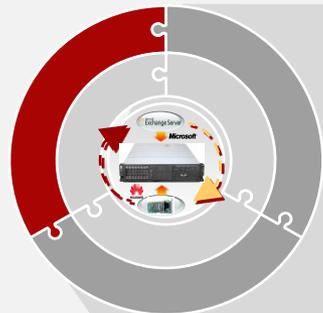
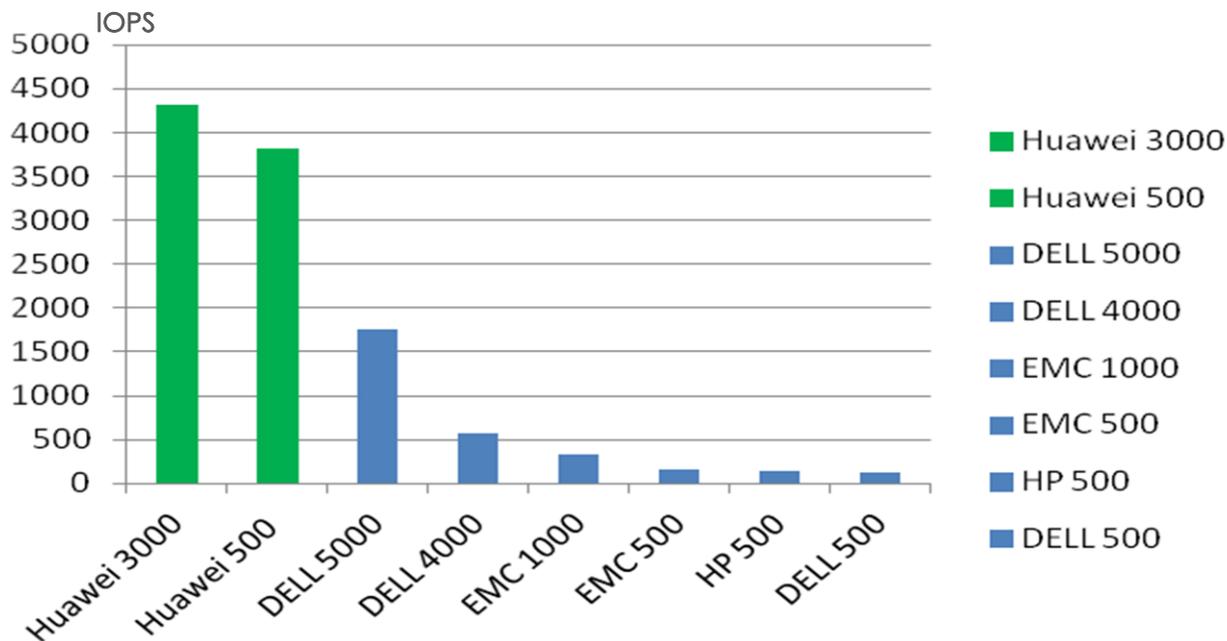


# 为Exchange 邮件服务器提供卓越平台



## 卓越的平台性能

- 支持500至3000个邮箱用户规模，用户数可平滑调节，在同类Exchange应用服务器性能测评(Jetstress)中，华为一体化方案比友商有大幅度的提升。



# PCIe SSD产品对比维度及ES3000分析

对比维度	对比建议	ES3000分析
<b>Flash颗粒</b>	MLC: 容量大, 性能一般, 价格较低, 可靠性一般 SLC: 容量小, 性能好, 价格高, 可靠性高 硬件是否同时兼容MLC、SLC	仅有MLC产品, 可根据市场需求, 兼容升级SLC
<b>容量</b>	相同尺寸情况下, 容量灵活组合扩展, 越大越好	800GB/1.2TB/2.4TB, 容量能够匹配各种业务应用需求
<b>尺寸</b>	主流尺寸: 全高半长 (FH-HL)、半高半长 (HH-HL或LP) 当前主流服务器很少预留全高全长尺寸PCIe槽位, 不建议用全高全长尺寸	全高半长卡, 兼容业界主流服务器安装
<b>总线接口</b>	业界主流采用 PCIe 2.0 x8接口	PCIe 2.0 x8
<b>性能</b>	4KB数据块 100%随机的读、写操作性能, 如IOPS、延迟、带宽	性能与主流厂商产品相当, 部分超越
<b>功耗</b>	相同容量和性能情况下, 功耗越低越好	性能好, 由于实现方案不同, 功耗稍高
<b>实现方案</b>	1、ASIC: 功耗低, 性能较高, 不支持定制化 2、FPGA: 性能高, 功耗稍高, 支持定制化 3、SATA-HBA: 类似SSD硬盘方案, 性能一般, 功耗高, 方案固定	FPGA实现方式, 后期规划采用ASIC方案
<b>可靠性</b>	ECC、磨损平衡、掉电保护等可靠性加固	内置ECC、RAID5、通道冗余、磨损平衡、掉电保护、TRIM技术等, 与业界主流一致
<b>管理维护</b>	提供管理工具, 且支持SNMP协议, 统一被机房管理系统管理	提供管理工具, 兼容SNMP协议
<b>配套软件及应用</b>	形成应用加速解决方案 (数据库、虚拟化、Cache等), 提供可选配套加速软件	当前缺少配套加速软件, 但已有规划
<b>成本</b>	按每GB价格计算	有优势

# 主流友商PCIe SSD产品对比

	FUSION-IO		STEC	OCZ		Huawei
型号	ioDrive2 Duo		Kronos™ Turbo PCIe SSA	Z-Drive R4 CM84/RM84	Z-Drive R4 CM88/RM88	ES3000
Flash颗粒	SLC	MLC	MLC	MLC	MLC	MLC
容量	1.2TB	2.4TB	1.95TB	300GB, 600GB, 1.2TB	800GB, 1.6TB, 3.2TB	800GB, 1.2TB, 2.4TB
尺寸	FH-HL	FH-HL	FH-HL	HH-HL	FH-3/4L	FH-HL
总线接口	Gen 2 x8	Gen 2 x8	Gen 2 x8	Gen 2 x8	Gen 2 x8	Gen 2 x8
读带宽	3.0GB/s	3.0GB/s	2GB/s	2GB/s	2.8GB/s	3.2GB/s
读IOPS	950k(512B)	892k(512B)	220k(4KB)	-	-	760K(4KB)
读延迟	47μs	68μs	-	-	-	49μs
写带宽	2.5GB/s	2.5GB/s	2GB/s	2GB/s	2.8GB/s	2.8 GB/s
写IOPS	935k(512B)	935k(512B)	200k(4KB)	250k(4kB) 160k(8kB)	410k(4kB) 275k(8kB)	240K(4KB)
写延迟	15μs	15μs	<50μs	-	-	8μs



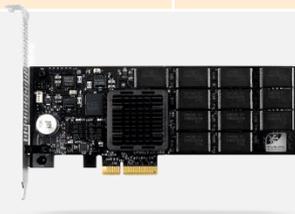
## HUAWEI ENTERPRISE ICT SOLUTIONS **A BETTER WAY**

**Copyright©2012 Huawei Technologies Co., Ltd. All Rights Reserved.**

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

# Fusion-io (ioDrive) 产品规格

型号	ioDrive				ioDrive Duo			
Flash颗粒	SLC		MLC		SLC		MLC	
容量	160GB	320GB	320GB	640GB	320GB	640GB	640GB	1.28TB
尺寸	HH-HL	HH-HL	HH-HL	HH-HL	FH-HL	FH-HL	FH-HL	FH-HL
PCIe接口	Gen 2 x4	Gen 2 x4	Gen 2 x4	Gen 2 x4	Gen 2 x4	Gen 2 x4	Gen 2 x4	Gen 2 x4
读带宽	770MB/s	770MB/s	735MB/s	750MB/s	1.5 GB/s	1.5 GB/s	1.5 GB/s	1.5 GB/s
读IOPS	140k(512B)	140k(512B)	100k(512B)	93k(512B)	261k(512B)	273k(512B)	196k(512B)	185k(512B)
读延迟	26μs	26μs	29μs	30μs	26μs	26μs	29μs	30μs
写带宽	750MB/s	790MB/s	510MB/s	550MB/s	1.5 GB/s	1.5 GB/s	1.0 GB/s	1.1 GB/s
写IOPS	123k(512B)	119k(512B)	67k(512B)	74k(512B)	262k(512B)	252k(512B)	285k(512B)	278k(512B)



参考: <http://www.fusionio.com/platforms/iodrive-duo/>

# Fusion-io (ioDrive2) 产品规格

型号	ioDrive2					ioDrive2 Duo	
Flash颗粒	SLC		MLC			SLC	MLC
容量	400GB	600GB	365GB	785GB	1.2TB	1.2TB	2.4TB
尺寸	HH-HL	HH-HL	HH-HL	HH-HL	HH-HL	FH-HL	FH-HL
PCIe接口	Gen 2 x4	Gen 2 x8	Gen 2 x8				
读带宽	1.4 GB/s	1.5 GB/s	910MB/s	1.5 GB/s	1.5 GB/s	3.0/s	3.0GB/s
读IOPS	480k(512B)	485k(512B)	415k(512B)	443k(512B)	445k(512B)	950k(512B)	892k(512B)
读延迟	47μs	47μs	68μs	68μs	68μs	47μs	68μs
写带宽	1.3 GB/s	1.3 GB/s	590MB/s	1.1 GB/s	1.3 GB/s	2.5GB/s	2.5GB/s
写IOPS	535k(512B)	535k(512B)	535k(512B)	535k(512B)	535k(512B)	935k(512B)	935k(512B)
写延迟	15μs	15μs	15μs	15μs	15μs	15μs	15μs
功耗	24W	24W	24W	24W	24W	55W	55W



参考: <http://www.fusionio.com/platforms/iodrive2-duo/>

# STEC产品规格

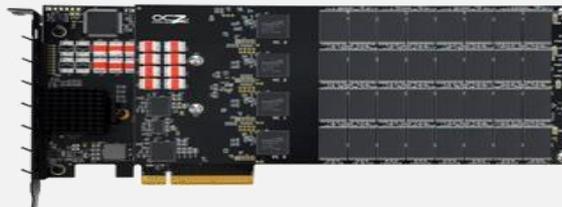
型号	Kronos™ PCIe SSA	Kronos™ Turbo PCIe SSA
Flash颗粒	MLC	MLC
容量	~980GB	~1.95TB
尺寸	HH-HL	FH-HL
PCIe接口	Gen 2 x4	Gen 2 x8
读带宽	1GB/s	2GB/s
读IOPS	110k	220k
写带宽	1GB/s	2GB/s
写IOPS	100k	200k
平均延迟	<50μs	<50μs



参考: <http://stec-inc.com/product/pciessa.php>

# OCZ产品规格

型号	Z-Drive R4 CM84/RM84	Z-Drive R4 CM88/RM88	Z-Drive R4 RS84	Z-Drive R4 RS88	Z-Drive R4 RM1616
Flash颗粒	MLC	MLC	SLC	SLC	MLC
容量	300GB, 600GB, 1.2TB	800GB, 1.6TB, 3.2TB	150GB, 300GB, 600GB	400GB, 800GB, 1.6TB	800GB, 1.6TB, 3.2TB, 6.4TB
尺寸	HH-HL	FH-3/4L	HH-HL	FH-3/4L	FH-3/4L
PCIe接口	Gen 2 x8	Gen 2 x8	Gen 2 x8	Gen 2 x8	Gen 2 x16
读带宽	2,000 MB/s	2,800 MB/s	2,000 MB/s	2,900 MB/s	6GB/s
写带宽	2,000 MB/s	2,800 MB/s	2,000 MB/s	2,900 MB/s	6GB/s
写IOPS	250k(4kB) 160k(8kB)	410k(4kB) 275k(8kB)	250k(4kB) 160k(8kB)	410k(4kB) 275k(8kB)	800k(4kB)
功耗	Idle: 14.5 W Active: 16 W	Idle: 23 W Active: 26 W	Idle: 14.5 W Active: 16 W	Idle: 23 W Active: 26 W	Idle: 23 W Active: 26 W



参考: <http://www.oczenterprise.com/interfaces/pci-express.html>

# Virident产品规格

型号	FlashMAX		FlashMAX II		
Flash颗粒	MLC	SLC	MLC	MLC	SLC
容量	1000, 1400GB,底板+模块	300, 400, 600, 800 GB,底板+模块	550GB	1100GB, 2200GB	1100GB
尺寸	HH-HL	HH-HL	HH-HL	HH-HL	HH-HL
PCIe接口	Gen 2 x8	Gen 2 x8	Gen 2 x8	Gen 2 x8	Gen 2 x8
读带宽	1.3GB/s(4kB)	1.4 GB/s (4KB)	1.6GB/s	2.7GB/s	2.7GB/s
读IOPS	325k IOPS (4KB) 1,000k IOPS (512 byte)	340k IOPS (4KB) 1,400k IOPS (512 byte)	160K(4KB)	325K(4KB)	525K(4KB)
读延迟	62μs	47μs	76μs(512B)	68μs(512B)	47μs(512B)
写带宽	600 MB/s	1.1GB/s	540MB/s	1GB/s	1.6GB/s
写延迟	19μs	16μs	16μs(512B)	18μs(512B)	18μs(512B)
功耗	24W	24W			



参考: <http://www.virident.com/products/flashmax/>

# Intel产品规格

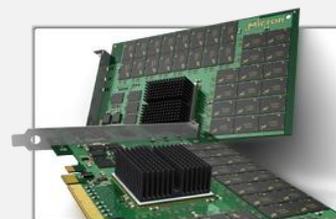
型号	Intel SSD 910	
Flash颗粒	MLC	
容量	400GB	800GB
尺寸	HHHL	HHHL
PCIe接口	Gen 2 x8	Gen 2 x8
读带宽	1GB/s	2GB/s
读IOPS	90k(4KB)	180k(4KB)
读延迟	65 $\mu$ s	65 $\mu$ s
写带宽	0.75GB/s	1GB/s
写IOPS	38k(4KB)	75k(4KB)
写延迟	65 $\mu$ s	65 $\mu$ s
功耗	Idle: 8 W Active: 25 W	Idle: 8 W Active: 25 W



<http://www.intel.cn/content/www/cn/zh/solid-state-drives/ssd-910-series-specification.html?wapkw=intel+ssd+910>

# Micron产品规格

型号	Micron RealSSD™ P320h
Flash颗粒	SLC
容量	350GB,700GB
尺寸	HH-HL
PCIe接口	Gen 2 x8
读带宽	3.2GB/s
读IOPS	785k(4K)
读延迟	<42μs(512B)
写带宽	1.9GB/s
写IOPS	205k(4K)
写延迟	<9μs
功耗	25W



参考：<http://www.micron.com/products/solid-state-storage/enterprise-pcie-ssd>