

TaoCloud | 大道云行  
MaxIO 缓存加速软件  
性能测试报告

Version 1.1



北京大道云行科技有限公司

—TaoCloud—

2015.10

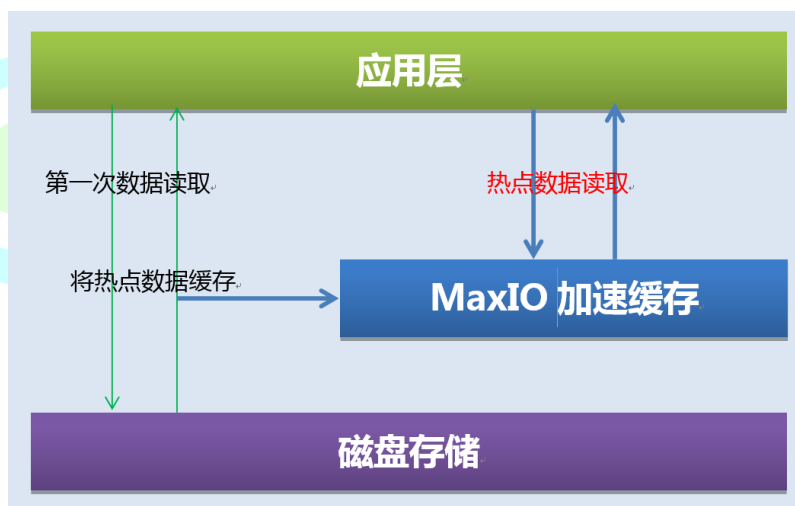
# 目录

TaoCloud   大道云行 MaxIO 缓存加速软件 性能测试报告 .....	1
目录 .....	2
一、MaxIO 简介 .....	3
二、测试环境 .....	4
1、硬件配置 .....	4
2、软件配置 .....	4
三、性能测试 .....	4
1、测试方法 .....	4
2、测试场景 .....	5
3、测试结果 .....	5
附录 1: MaxIO 操作简介 .....	7
1、    MaxIO Cache 创建 .....	7
2、    MaxIO Cache 删除 .....	7
3、    MaxIO Cache 信息 .....	7
附录 2: FIO 测试配置 .....	8
1、    顺序写—seq-100w0r.fio .....	8
2、    顺序读—seq-0w100r.fio .....	9
3、    随机写—rand-100w0r.fio .....	10
4、    随机读—rand-0w100r.fio .....	11
5、    随机混合—rand-25w75r.fio .....	12

## 一、MaxIO 简介

TaoCloud MaxIO 是一款以 SSD 为核心硬件加速基础设计的智能缓存加速软件，可在保持 SSD 长效使用的前提下，充分利用 SSD 的 IO 特点和性能优势，针对 HDD 存储卷进行高效加速。

MaxIO可以部署在任意X86硬件上，包括服务器，软件定义存储系统，因此可以在存储层和计算层灵活实现加速，且仅占用极少的系统CPU和内存资源，能够为更多的应用提供加速效果。MaxIO采用透明的加速配置，可在线增加或删除加速模块，配置时，应用无需任何配置更改，因此对企业应用场景不产生任何干扰。MaxIO具有友好的硬件兼容性，支持各种厂商的SATA SSD或PCIe SSD。



MaxIO缓存加速原理

MaxIO采用了独特且高效的缓存技术来加速企业级存储应用，对IO进行智能分析，根据加速设备SSD的大小，将更多更常用的热点数据放置于高速SSD介质中，减少SSD的存取次数，延长SSD的使用寿命，并缩短应用等待数据的时间。

MaxIO支持多种灵活的加速模式，针对不同的业务场景，可发挥不同的加速效果，包括透写/回写/只读模式。

## 二、测试环境

### 1、硬件配置

硬件项目	配置
X86 服务器	双路 E5 2630v2 CPU 2.6GHz; 64GB 内存;
HDD	西数 7200rpm 2TB SATA
SATA SSD	Intel DC S3710, 800GB
PCIe SSD	Comay BladeDrive E24 800GB

### 2、软件配置

软件项目	配置
Linux OS	CentOS 6.6 x86_64 (2.6.32-504)
MaxIO	maxio-1.1.1-260.x86_64
FIO	fio-2.0.13-1.el6.x86_64

## 三、性能测试

### 1、测试方法

性能是评估一个存储产品最为关键的指标，包括 IOPS、带宽和延迟，这里采用 FIO 进行测试分析。相关的 FIO 测试配置，详见附录。

FIO 是一个 I/O 标准测试和硬件压力验证工具，它支持 13 种不同类型的 I/O 引擎，支持块设备和文件系统测试，广泛用于标准测试、QA、验证测试等，支持 Linux, Windows 等操作系统。

采用 FIO 测试 HDD 和 SSD 的基准数据，然后再测试 SSD 对 HDD 进行缓存加速后的性能，包括顺序读写、随机读写和混合读写性能指标。通过对比测试验证 MaxIO 的性能加速效果。

## 2、测试场景

测试场景	场景描述	测试结果
HDD 单盘基准	对单块 HDD 进行基准性能测试，分别测试 256KB/4KB 的顺序和随机读写性能。FIO 测试配置见附录。	
SATA SSD 单盘基准	对单块 SATA SSD 进行基准性能测试，分别测试 256KB/4KB 的顺序和随机读写性能。FIO 测试配置见附录。	
SATA SSD 缓存加速	对 SATA SSD 加速后的 HDD 进行性能测试，分别测试 256KB/4KB 的顺序和随机读写性能。FIO 测试配置见附录。	
PCIe SSD 单盘基准	对单块 PCIe SSD 进行基准性能测试，分别测试 256KB/4KB 的顺序和随机读写性能。FIO 测试配置见附录。	
PCIe SSD 缓存加速	对 PCIe SSD 加速后的 HDD 进行性能测试，分别测试 256KB/4KB 的顺序和随机读写性能。FIO 测试配置见附录。	

## 3、测试结果

MaxIO 性能与硬件配置密切相关，理论上硬件配置越高性能越好，本次测试的硬件配置详见硬件环境描述，软件设置采用系统缺省配置。为了保证性能测试结果的客观性，作如下测试前置条件：

- (1) FIO 策略：100G, direct, numjobs=4, iodepth=16(顺序)/64(随机)
- (2) MaxIO 策略：LRU 替换算法, writeback, block\_size=4096
- (3) 测试前采用 FIO 将测试磁盘完全顺序写一遍；
- (4) 每组 FIO 的测试时长为 300 秒；
- (5) 每次测试前清空缓存，保证本次结果不受上次测试影响；

在上述测试前提下，得到如下 HDD/SSD 基准性能和 MaxIO 缓存加速性能测试结果。其中，MaxIO 缓存加速性能数据同时给出了命中率 HitRate，这个直接关系加速效果。

	顺序写 (MB/s)	顺序读 (MB/s)	随机写 (IOPS)	随机读 (IOPS)	随机读写 (W:R=1:3)
HDD	107	113	238	233	N/A
SATA SSD	268	279	38092	49687	49870
MaxIO 加速 (SATA SSD)	270 (Hit: 70%)	298 (Hit: 94%)	38279 (Hit: 71%)	49721 (Hit: 96%)	48048 (Hit: 99/72%)
PCIe SSD	1777	2012	148768	164449	156077
MaxIO 加速 (PCIe SSD)	1454 (Hit: 99%)	1700 (Hit: 99%)	101667 (Hit: 96%)	151199 (Hit: 99%)	139588 (Hit: 99/96%)

MaxIO 缓存加速性能测试结果

从上表的性能测试数据来看，MaxIO 对 SATA SSD 和 PCIe SSD 都可以获得非常好的性能加速效果，包括顺序写(SeqWrite)、顺序读(SeqRead)、随机写(RandWrite)、随机读(RandRead)、随机混合读写(RandRW)性能。其中，MaxIO 对 SATA SSD 分别获得了 100%、106%、100%、100%、96% 的加速比；MaxIO 对 PCIe SSD 分别获得了 82%、84%、68%、92%、89% 的加速比。可见，在对 SSD 进行预热后，可以获得相当高的读写缓存命中率，从而获得非常可观的性能加速效果，显著提升普通磁盘的性能，甚至达到裸 SSD 的性能。

值得注意的是，MaxIO 缓存加速效果与应用 I/O 模式密切相关。通常情况下，SSD 与 HDD 的加速容量配比为 10~20%，本次配置的 SSD 容量较大，而且测试数据大小为 100G，因此效果良好。如果 SSD 容量较小，则 SSD 缓存超过设定阈值后，MaxIO 读写同时会进行 Flush 数据刷回 HDD 操作，加速效率会受到较大的影响。

正常应用场景下，开始时 MaxIO 缓存没有数据，命中率低，加速效果会欠佳。随着 I/O 读写不断加热，MaxIO 缓存数据不断增加，命中率提高，逐渐达到较佳的加速效果。另外，MaxIO 对随机读写加速效果显著，系统会智能判断顺序 I/O 模式并直接写 HDD，因此在一些场景下顺序读写效果会表现不佳。



## 附录 1: MaxIO 操作简介

### 1、MaxIO Cache 创建

```
maxio create [-h] -s SSD -p HDD [-r {rand,fifo,lru}] [-t {wb,wt,ro}]  
[-b {2048, 4096, 8192}] [-f] cache_name
```

参数说明:

- s SSD 必选参数, 系统中SSD分区或设备, 可使用lsblk查看当前系统中所有设备;
- p HDD 必选参数, 需要加速的设备或分区, 可使用lsblk查看当前系统中所有设备;
- r {rand,fifo,lru} 可选参数, 缓存调度策略, 默认为lru;
- t {wb,wt,ro} 可选参数, 缓存模式, 默认为wt, 即透写;
- b {2048,4096,8192} 可选参数, 缓存的块大小, 默认为 4096;

例如:

```
maxio create -s /dev/sdb -p /dev/sdc cache1
```

### 2、MaxIO Cache 删除

```
maxio delete cache_name
```

特别说明:

删除 Cache 时, 需要把 SSD 中缓存的数据写回低速 HDD, 尤其是大容量 SSD, 时间会相对比较长, 请特别注意。如果数据未写回完成, 可能导致数据不一致。

例如:

```
maxio delete cache1
```

### 3、MaxIO Cache 信息

```
maxio show cache_name
```

说明: 显示 cache 设备信息, 以及读写的缓存命中率。

例如:

```
maxio show cache1
```

## 附录 2: FIO 测试配置

### 1、顺序写—seq-100w0r.fio

```
[global]
ioengine=libaio
direct=1
thread=1
norandommap=1
randrepeat=0
ramp_time=0
filename=/dev/XXX (请修改设备名)
```

```
[seqwrite]
stonewall
group_reporting
time_based
size=100g
bs=256k
rw=write
numjobs=4
iodepth=16
runtime=600
```





## 2、顺序读—seq-0w100r.fio

```
[global]
ioengine=libaio
direct=1
thread=1
norandommap=1
randrepeat=0
ramp_time=0
filename=/dev/XXX (请修改设备名)
```

```
[seqread]
stonewall
group_reporting
time_based
size=100g
bs=256k
rw=read
numjobs=4
iodepth=16
runtime=600
```



### 3、随机写—rand-100w0r.fio

```
[global]
ioengine=libaio
direct=1
thread=1
norandommap=1
randrepeat=0
ramp_time=0
filename=/dev/XXX (请修改设备名)
```

```
[randwrite]
stonewall
group_reporting
time_based
size=100g
bs=4k
ba=4k
rw=randwrite
numjobs=4
iodepth=64
runtime=600
```



#### 4、随机读—rand-0w100r.fio

```
[global]
ioengine=libaio
direct=1
thread=1
norandommap=1
randrepeat=0
ramp_time=0
filename=/dev/XXX (请修改设备名)
```

```
[randread]
stonewall
group_reporting
time_based
size=100g
bs=4k
ba=4k
rw=randread
numjobs=4
iodepth=64
runtime=600
```



## 5、随机混合—rand-25w75r.fio

```
[global]
ioengine=libaio
direct=1
thread=1
norandommap=1
randrepeat=0
ramp_time=0
filename=/dev/XXX (请修改设备名)
```

```
[randrw]
stonewall
group_reporting
time_based
size=100g
bs=4k
ba=4k
rw=randrw
rwmixread=75
numjobs=4
iodepth=64
runtime=600
```