

如何避免生产环境性能瓶颈

MySQL压力测试经验

叶金荣，ORACLE MySQL ACE

<http://imysql.com>， 微信公众号：[MySQL中文网](#)， weibo：[@yejinrong](#)

2013-11-17

- 叶金荣，网络常用ID：yejr
- Oracle MySQL ACE
- 国内最早的MySQL推广者
- 2006年创办国内首个MySQL专业技术网站
<http://imysql.com>
- 资深MySQL专家，10余年MySQL经验，擅长MySQL性能优化、架构设计、故障排查



作为DBA，碰到下面情况如何应对

- 开发新项目时，如何评估数据库性能
- 新系统上线前，如何预估数据库负载
- 采购新设备，如何评估其性能
- 更换数据库版本，如何评估其性能

大纲

- 关键指标
- 测试工具
- 性能基准测试
- 可靠性测试
 - 模拟意外事件
 - 长期极限高压
 - 模拟恶劣环境
- 测试报告
- 附件下载

测试环境

技术指标	测试环境A	测试环境B
机型	DELL PE R720 (2U PC Server)	DELL PE R710 (2U PC Server)
CPU	Xeon E5-2620(6核 , 12线程 , 2.0GHz, L3 15MB) * 2	Xeon E5620(4核 , 8线程 , 2.4GHz , L3 12MB) * 2
内存	32G(4G * 8)	32G(4G * 8)
阵列卡及设置	PERC H710 , 512MB , BBU(FW : 12.10.1-0001) , RAID 1+0 FORCE WB	PERC H700 , 512MB , BBU(FW : 12.10.1-0001) , RAID 1+0 FORCE WB
硬盘	15K RPM 300G SAS * 8	15K RPM 300G SAS * 6
网卡	Intel 1GbE	Broadcom 1GbE
操作系统	RHEL 6.4	RHEL 6.4
文件系统	xfs/ext4	xfs/ext4
MySQL版本	5.5.34	5.6.14

关键指标

- 关键指标
 - CPU : %wait , %user , %sys
 - 内存 : 只内存读写 , 不产生swap
 - 磁盘 : IOPS、响应时长
 - 数据库 : TPS/TpmC、吞吐量 (QPS) 、响应时长、InnoDB
- 注意事项
 - 本地加压
 - 数据量小
 - 压力时间过短
 - 压测模式太少 (读写比例变化、并发数变化)
 - 压力过大或过小

性能基准测试 – 关注点

- CPU
 - 简单素数计算、复杂浮点计算
 - thread分配
 - mutex性能
- 关注信息
 - 计算能力
 - 多线程并发
 - mutex管理

常地，随着计算机工艺的发展，大部分应用下，CPU几乎不再成为瓶颈。。一般只关注不同品牌厂商、不同工艺、不同主频下的不同表现，其他情况下，无需太多关注。

性能基准测试 – 关注点

- 内存
 - 总带宽
 - 读写效率
 - CPU对内存的管理分配
- 关注信息
 - 读写效率

通常，内存越大越好，可有效减少磁盘物理IO。一般只关注不同品牌厂商、不同工艺、不同主频下的不同表现，其他情况下，无需太多关注。

性能基准测试 – 关注点

- 磁盘IOPS
 - 读写效率
 - 随机写性能
- 关注信息
 - 不同内核、文件系统下的IOPS
 - 不同阵列级别、条带场景下的IOPS
 - 和内存结合时的IOPS性能拐点
 - 达到同样IOPS情景下，IO利用率差别
 - 跑满IO以及IO压力较轻情景下，各自IO利用率差别
 - `sar -d : tps、svctm、%util`

通常，磁盘物理IO是这个计算机体系里最容易成为瓶颈的环节，也是最难优化的，因此最需要关注。随着SSD、Fusion-IO出现，磁盘IOPS获得了巨大提升，和内存相比，差距在不断缩小。

性能基准测试 – 关注点

- OLTP/OLAP
 - TPS、QPS、
 - 响应延迟/分析、吞吐效率
- 关注信息
 - 不同内核、文件系统下的TPS
 - 不同阵列级别、条带场景下的TPS
 - 和内存结合时的TPS性能拐点
 - 达到同样TPS情景下，IO负载差别
 - TPS满负荷、低负荷情景下，各自IO利用率差别
 - `sar -d : tps、svctm、%util`

通常，数据总量超过物理内存后，OLTP的TPS性能和磁盘IOPS成正比关系。因此，只有不断提高内存，减少物理IO，并且不断提升IOPS性能。

性能基准测试 – 关键因素

- 硬件
 - CPU、内存、阵列卡(BBU、CACHE、条带、读写策略)、硬盘
- 系统
 - 内核参数、文件系统、IO调度器
- 文件
 - 块大小、访问模式
- 其他
 - 网络环境对over lan请求测试影响大
 - 每完成一轮测试后要净化环境
 - 每轮测试一般至少持续1小时

测试工具

- sysbench
 - Primarily for MySQL OLTP benchmarking , By MySQL AB
 - cpu、 threads、 mutex、 memory、 fileio、 oltp
- iozone
 - filesystem benchmark tool
- tpcc-mysql
 - Primarily for MySQL OLTP benchmarking , By Percona
- tpch
 - Primarily for OLAP benchmarking
- 其他
 - OLTP : mysqlslap、 sql-bench
 - IOPS : bonnie、 orion、 iometer
 - 综合 : stress

测试工具 - sysbench

- 安装
 - `./configure --with-mysql-includes=path --with-mysql-libs=path && make && make install`
- 支持其他数据库
 - `with-pgsql`
 - `with-oracle`
- 运行
 - `sysbench --test=[mode] [other_options] prepare`
 - `sysbench --test=[mode] [other_options] run`
 - `sysbench --test=[mode] [other_options] cleanup`

测试工具 - sysbench

- 通用基准
 - 最大请求数：5,000,000
 - 并发线程数：8 ~ 512
- 基准 - cpu
 - cpu-max-prime=500,000
- 基准 - threads
 - thread-yields=5,000,000
 - thread-locks=8~512

测试工具 - sysbench

- 基准 - mutex
 - mutex-num=5,000,000
 - mutex-loops=100,000
 - mutex-locks=100,000
- 基准 - fileio
 - file-num=100
 - file-total-size=物理内存 ~ 物理内存的8倍
 - mode=seqwr/seqrewr/seqrd/rndrd/rndwr/rndrw

测试工具 - sysbench

- 基准 - memory
 - mode=complex
 - engine=innodb
 - oltp-table-size=100,000,000
- 基准 - OLTP
 - mode=complex
 - engine=innodb
 - oltp-table-size=100,000,000

测试工具 - sysbench

- fileio mode

seqwr	顺序写
seqrewr	顺序重写
seqrd	顺序读
rndrd	随机读
rndwr	随机写
rndrw	随机读写

- OLTP mode

nontrx	无事务
simple	简单请求
complex	复杂请求，有事务

测试工具 - iotzone

- 安装
 - make linux-AMD64
- 运行
 - iotzone -R -E -s 3200M -l 10 -r 4k

测试工具 - iotzone

- 基准 - fileio
 - file-block-size=4k 8k 16k 32k 64k
 - file-num=1 2 4 8 16
 - file-size=1G 2G 4G 8G 16G
 - mode=seqwr/seqrewr/seqrd/rndrd/rndwr/rndrw

测试工具 - tpcc mysql

- 安装

- 下载 bzr branch lp:~percona-dev/perconatools/tpcc-mysql
- 直接make即可
- create_table.sql - 创建数据表
- add_fkey_idx.sql - 创建索引及外键

- 初始化加载数据

- tpcc_load db_host db_name db_user db_passwd
db_warehouse_num
- 例如：tpcc_load localhost tpcc1000 user passwd 1000

- 运行OLTP测试

- ./tpcc_start -h localhost -d tpcc1000 -u root -p 'xx' -w 1000
-c 32 -r 120 -l 3600 -f ./tpcc_mysql_20120314

测试工具 - tpcc mysql

- 基准 - OLTP
 - warehouse = 1000
 - max connection = 8 ~ 512
 - warm up = 120(s)
 - run time/duration = 3600(s)

测试工具 – tpch

- 安装

- cp makefile.suite makefile
- 编辑 makefile 文件

```
CC = gcc
```

```
DATABASE = MYSQL
```

```
MACHINE = LINUX
```

```
WORKLOAD = TPCH
```

- 编辑tpcd.h , 增加宏定义

```
#ifdef MYSQL
```

```
#define GEN_QUERY_PLAN ""
```

```
#define START_TRAN "START TRANSACTION"
```

```
#define END_TRAN "COMMIT"
```

```
#define SET_OUTPUT ""
```

```
#define SET_ROWCOUNT "limit %d;\n"
```

```
#define SET_DBASE "use %s;\n"
```

```
#endif
```

测试工具 – tpch

- 初始化

- 初始化测试表数据： `./dbgen -s 100`
- 生成测试数据： `mysql -f tpch < dss.ddl`
- 默认的初始化模式无主键、无索引
- `LOAD DATA INFILE`导入数据
- 注意`max_binlog_cache_size`限制，需要切分文件导入
- 执行修改主键/外键/额外索引脚本
- 数据表名全部改成小写，适应TPC-H测试SQL脚本
- 运行qgen生成测试SQL
- 修改部分SQL语句
- 拆分完成测试SQL脚本成23个测试SQL

- 运行

- 执行23个测试脚本，记录运行时长

测试工具 – tpch

- 基准 - tpch
 - warehouse = 100
 - 单进程
 - tpch侧重OLAP模型，而MySQL并不适合OLAP，因此warehouse设定较小

测试报告

- cpu - sysbench

线程数	最大请求数	计算最大素数	时间(s)	请求统计		
				最快	最慢	平均
32	500,000	5,000,000	275258.028	8.788	27.615	17.616
32	500,000	500,000	10982.417	0.332	2.370	0.703
64	500,000	500,000	10981.551	0.351	4.853	1.405
64	5,000,000	500,000	109815.115	0.351	6.418	1.406

测试报告

- memory - sysbench

读内存									
线程数	最大请求数	内存总大小 (GB)	内存块大小(K)	读性能 (ops/sec)	读速度 (MB/sec)	读耗时(s)	请求统计		
							最快	最慢	平均
4	50000	4	4	531823.9500	2077.4400	1.9717	0.0000	0.0007	0.0000
4	50000	4	8	328783.8900	2568.6200	1.5946	0.0000	0.0005	0.0000
4	50000	8	4	543861.0100	2124.4600	3.8560	0.0000	0.0009	0.0000
4	50000	8	8	350874.5600	2741.2100	2.9885	0.0000	0.0009	0.0000

写内存									
线程数	最大请求数	内存总大小 (GB)	内存块大小(K)	写性能 (ops/sec)	写速度 (MB/sec)	写耗时(s)	请求统计		
							最快	最慢	平均
4	50000	4	4	58778.210	229.600	17.840	0.0001	0.0012	0.0001
4	50000	4	8	30567.630	238.810	17.152	0.0001	0.0012	0.0001
4	50000	8	4	59182.720	231.180	35.435	0.0001	0.0107	0.0001
4	50000	8	8	29886.260	233.490	35.086	0.0001	0.0056	0.0001

测试报告

- fileio - sysbench

线程数	测试模式	最大请求数	文件个数	文件总大小(GB)	文件块大小(KB)	传输速度(Mb/sec)	执行总时间(s)	请求统计		
								最快	最慢	平均
4	rndrw	1,000	10	4	4	5.2507	0.7440	0.0000	0.0776	0.0002
4	rndrw	1,000	10	4	8	8.8667	0.8811	0.0000	0.0043	0.0001
4	rndrw	10,000	10	4	8	8.4942	9.1975	0.0000	0.2442	0.0005

测试报告

- fileio - iozone

线程数	文件总大小 (GB)	文件块大小 (KB)	Initial_write (MB/ s)	Rewrite (MB/ s)	Read (MB/ s)	Re-read (MB/ s)	Reverse_Read (MB/ s)	Stride_read (MB/ s)
10	32	4	700.701	752.705	12,519.545	13,211.167	11,634.753	10,734.569
10	32	8	712.641	716.450	13,446.036	13,136.648	10,336.330	10,510.443
10	32	16	719.171	756.499	12,610.633	13,090.735	9,503.751	9,728.941
10	32	32	714.208	730.096	12,931.627	12,248.437	10,554.508	10,989.790
10	64	4	496.580	490.825	684.361	5,065.913	4,016.659	3,136.843
10	64	8	495.682	494.908	246.391	245.638	79.939	19.264
10	64	16	493.391	492.137	249.678	257.989	115.421	29.950

Random_read (MB/ s)	Mixed_workload (MB/ s)	Random_write (MB/ s)	Pwrite (MB/ s)	Pread (MB/ s)	Fwrite (MB/ s)	Fread (MB/ s)	耗时 (秒)
9,563.427	7,053.637	55.844	665.776	1,102.762	737.492	5,455.919	1,560.249
9,406.446	7,165.425	95.712	670.443	1,116.525	746.309	6,844.481	1,096.531
8,788.983	1,609.540	154.505	706.784	1,557.716	724.744	5,742.960	839.051
10,715.645	6,718.890	228.905	702.704	884.447	792.748	3,814.560	710.959
2,487.435	58.087	22.711	495.547	595.622	501.170	279.480	5,816.913
11.154	97.538	40.192	495.200	248.907	496.403	268.686	10,065.453
21.740	153.103	67.479	494.486	252.050	500.590	263.018	6,541.998

测试报告

- oltp - sysbench

线程数	第一轮 (tps)	第二轮 (tps)	第三轮 (tps)	均值 (tps)
16	1266.867	1267.683	1267.275	1267.275
32	1744.833	1744.350	1744.592	1744.592
64	1891.250	1884.600	1887.925	1887.925
128	1938.483	1923.983	1931.233	1931.233
256	1920.350	1913.017	1916.684	1916.684

测试报告

- oltp - tpcc

线程数	第一轮 (TpmC)	第二轮 (TpmC)	第三轮 (TpmC)	均值 (TpmC)
16	1266.867	1267.683	1267.275	1267.275
32	1744.833	1744.350	1744.592	1744.592
64	1891.250	1884.600	1887.925	1887.925
128	1938.483	1923.983	1931.233	1931.233
256	1920.350	1913.017	1916.684	1916.684

测试报告

- olap – tpch

查询	第一轮	第二轮	第三轮	均值 (s)
Q1	2701.215	2697.523	2704.187	2700.975
Q2	416.314	418.774	417.474	417.521
Q3	3667.766	3755.572	3692.900	3705.413
Q4	105.804	105.493	102.411	104.569
Q5	2525.357	2536.164	2530.909	2530.810
Q6	2202.317	2194.628	2201.490	2199.478
Q7	2829.224	2860.434	2890.436	2860.031
Q8	3335.373	3383.118	3417.130	3378.540

可靠性测试

- 模拟意外事件
 - 断电(硬件冷重启)
 - RESET(硬件热重启)
 - 阵列卡掉线
 - 磁盘掉线
 - REBOOT(系统重启)
 - 正常关闭服务(kill -TERM)
 - 异常关闭服务(kill -9)
 - 磁盘空间满
 - 删除文件
 - 破坏性修改已打开文件
 - ...

可靠性测试

- 长期极限高压
 - 持续数小时、数天、数周运行高负载计算、IO任务
 - 考验服务器在高压下的性能波动情况
 - 考验硬件设备在高压下的稳定性表现
- 模拟恶劣环境
 - 供电不稳
 - 通风冷却不好
 - 湿气大、灰尘多

附件

- 整合sysbench测试脚本 [\[下载\]](#)
- 整合iozone测试脚本 [\[下载\]](#)
- 整合tpcc-mysql测试脚本 [\[下载\]](#)
- 整合tpch测试脚本 [\[下载\]](#)
- 汇总下载 [\[下载\]](#)
- 测试结果表格模板 [\[下载\]](#)